Contents

Tender Documents Procurer's version 28/04/2014

Procuring organization

Procurement Linköpings universitet AC2018/NWP18 for NSC in Linköping Hans Kempe LIU_DIARY_NUMBER

Legend: The text is included in the advert The text will be published in the contract catalogue The text/question contains requirements to be met The text

contains confidential information The text will be part of the contract The question is weighted and is part of the part of the procurement The text is included in the qualification The question is answered by the buyer

1 Invitation

1.1 Administrative regulations

1.1.1 General information

Linköping University (LiU) is a research-based university with excellence in education. LiU is a multi-faculty university where research and education are equally important. LiU has 27,000 undergraduate students, 4,000 staff and faculty members and a turnover of 3,400 million SEK. LiU hosts one of the major HPC (High Performance Computing) centres in Sweden named National Supercomputer Centre (NSC, http://www.nsc.liu.se).

NSC is a national supercomputing centre within SNIC (Swedish National Infrastructure for Computing). Current staff is about 35 people. NSC has served the Swedish academic community since 1989 as a provider of leading-edge supercomputing resources and storage services to members of academic institutions throughout Sweden and to NSC's partners SMHI (Swedish Meteorological and Hydrological Institute) and SAAB Defence and Security. NSC owns and operates a number of large-scale compute and storage resources. NSC also offers support and help to its users to enable best possible performance and efficient usage of the resources.

NSC is and has been active in several European projects such as ENACTS, HPC4U, IS-ENES1, IS-ENES2, CLIPC, EGI, PRACE 1-ip through PRACE 5-ip. NSC contributes to Swedish HPC research by its membership in the e-infrastructure organisation SeRC (Swedish e-Science Research Centre)

1.1.2 Scope

The purpose of this procurement is to purchase a HPC, have it installed at NSC's datacenter in Linköping, and sign a service and support contract for four (4) years with possibility to prolong with one (1) plus one (1) year.

1.1.3 Definitions

See Appendix 4 Definitions and Tables.

1.1.4 Contracting entity

Linköping University
SE-581 83 Linköping
Sweden

1.1.5 Contact

Katja Ekström, Procurement Officer
e-mail: katja.ekstrom@liu.se

1.1.6 Procurement procedure

1.1.6.1 Procedure This is a negotiated procurement procedure, in accordance with the Swedish law of public procurements (2016:1145). Tenders may be awarded without prior negotiation.

The procurement will be managed through http://www.opic.com, which is an online procurement system. The system contains tools to facilitate the procurement process for the contracting authority as well as for tenderers. Questions regarding the procurement system Visma OPIC TendSign should be submitted by mail to tendsignsupport@opic.com. Support can also be given by phone: +46 13 47 47 570

1.1.6.2 Type of contract LiU intends to award a contract to one supplier for the entire scope of the procurement.

1.1.6.3 Questions and answers If the contract document is considered to be unclear, it is important that the contact person for these matters be informed as early as possible to avoid any misunderstandings.

Enquiries during the tendering period must be made using the "Question and answer" function in www.opic.com. Linköping University will answer the questions through the same system.

All questions and answers shall be conducted in English.

The "Question and answer" function is available through the Tender notice under "Questions and Answers". Those who have not registered in the system are obliged to monitor any questions and answers during the procurement procedure by themselves.

Questions must be submitted at the latest by QA_DEADLINE, to allow enough time to communicate answers.

1.1.6.4 Additions / amendments Complementary additions/amendments, if any, will be announced through "Questions and Answers" in TendSign.

1.1.6.5 Cost of Tender The costs involved in preparing, submitting, and follow up the Tender and the procurement process will not be reimbursed.

1.1.7 The Tender - formal requirements

1.1.7.1 Validity of the Tender The Tender is valid provided that it is:

- signed by an authorised representative of the Tendering company,
- received no later than TENDER_DEADLINE. The Tenderer is solely responsible for the timely arrival of the Tender. A Tender that arrives later than stipulated will not be considered,
- valid until TENDER_VALID_UNTIL_DATE, which explicitly must be confirmed in the Tender.

Is the Tender valid until TENDER_VALID_UNTIL_DATE? (Yes/No answer)

1.1.7.2 Formality of the Tender Formal requirements for the Tenders are as follows:

- written in the Swedish or English language,
- Appendix 1 TenderForm must be printed, signed and enclosed with the Tender, e.g. as a PDF document.

1.1.7.3 Submission of the Tender The Tender should preferably be submitted through the tendering system Tendsign.com in electronic format. Please note that any included attachments should be of a format which is readily readable, like .pdf,.xls, .doc.

If a supplier cannot submit a Tender through the above mentioned system it can either be sent by post, by express, or be delivered personally at the address stated below. In case of submittance of Tender in hard copy then the following applies:

- The Tender must be submitted in a closed envelope/package marked with "Tender" and reference number "LIU_DIARY_NUMBER". Please note that tenders submitted by delivery firms often has to be marked on both inner and outer package,
- The Tender must be submitted in one (1) original and two (2) complete copies,

- and an electronic version on a CD/memory stick should also be provided.

Please note that the contracting authority will not accept any Tenders sent by fax or e-mail because the confidentiality cannot be assured.

Visiting, mailing, and express delivery address:

Linköping University
Registrator, Building Origo
SE-581 83 Linköping
Sweden

Office hours are Monday – Friday 09 AM – 3 PM. There is a mailbox outside the Origo building where Tenders can be delivered, but it is very limited in opening size (27,5 cm x 3,5 cm). Keep in mind that Tenders are submitted at the risk of the Tenderer, and to ensure that the Tender can be submitted correctly, the office hours should be taken into consideration.

1.1.7.4 Disposition of the Tender To ensure equal treatment and to facilitate assessment of Tenders, the Tenderer is requested to follow the order and numbering of this document.

If no statement is made with respect to a requirement, LiU is entitled to interpret this as an unfulfilled requirement.

1.1.7.5 System documentation in the Tender In the Tender the following System documentation is mandatory:

Description of the tendered solution:

- Brand and models of main components,
- Description of processor architecture,
- Network topology for communication networks,
- Rack layout including placement of components,
- Electrical connections,
- Physical size and weight of each rack.

Are these descriptions included in the Tender? (Yes/No answer)
..... yes ☐ no ☐

1.1.8 Evaluation process

1.1.8.1 Qualification of the Tenderer

Before the evaluation process starts, the formal requirements are checked in accordance with 1.1.7. If any of those are not fulfilled, the Tender in question may not be taken into consideration.

1.1.8.2 Examination

The requirements indicated as "must" in this document are checked. Only Tenders that fulfill all these requirements will be evaluated further.

1.1.8.3 Evaluation

The evaluation of the Tenders is based on the committed AC2018/NWP18 system performance T_tot described in subsection 4.4.1.

The Total cost for the AC2018/NWP18 system, including hardware, installation, service & support, and lifetime running cost must not exceed TCO_BUDGET SEK as detailed in "Appendix 3: AC2018/NWP18_summary".

Evaluation criteria

- The Tender with the highest committed AC2018/NWP18 system performance T_tot (with one (1) decimal), calculated in "Appendix 3 AC2018/NWP18_summary", will be awarded the contract. Highest performance should be taken to mean lowest T_tot in this context.
- If T_tot are evaluated as equal in Tenders the lowest Contract sum will be awarded the contract.
- If Contract sum and T_tot are evaluated as equal in Tenders, the fastest total run time of the VASP benchmarks will be awarded the contract. FIXME
- If T_tot are evaluated as equal in Tenders and the Contract sum are equal the contract will be awarded by drawing lots.

1.1.9 Commercial secrecy

In accordance with the Official Secrets Act (Offentlighets- och sekretesslag (2009:400)) all data pertaining to a procurement matter is subject to secrecy until an agreement has been signed or procurement has otherwise been concluded.

Any data mentioned may be subject to secrecy even after the aforementioned time. Note, however that only in exceptional cases data and prices according to the evaluation of the Tender is considered to be of such nature that they may be held secret for commercial reasons. With respect to commercial secrecy for the protection of a Tenderer's data, the requirement is that the data in question refers to the Tenderer's business- or service conditions and that there is a specific reason to presume that the Tenderer will suffer damages if the data is disclosed.

If a Tenderer considers any data submitted in connection with this procurement matter to fulfill the aforementioned requirements for commercial secrecy, the Tenderer must submit a request for commercial secrecy, in writing, with clarification concerning the data referred to and what damage the Tenderer would suffer if the data were disclosed.

The decision whether or not the data submitted by the Tenderer fulfill the requirements for commercial secrecy will be made by LiU. This decision could be challenged by a court of law, therefore no guarantees of secrecy can be given by LiU.

1.1.10 Notification of decision

Prior to signing the public contract, notification of LiU´s decision in this procurement matter will be made in TendSign.

An appeal for reconsideration of this decision can be made to Förvaltningsrätten (the Administrative Court) in Linköping. Such an appeal has to be sent within 10 days from notification.

Address:
Förvaltningsrätten i Linköping
Box 406
SE-581 04 Linköping
Sweden

E-mail: forvaltningsrattenilinkoping@dom.se

1.1.11 Agreement

Please note that there is no legally binding agreement until a public contract has been signed by both parties.

1.1.12 Site inspection

If the Tenderer wishes to perform a site inspection before submitting a tender either send a request through TendSign or send a mail to LiU's contact person for the procurement.

1.2 Documents

1.2.1 Provided by LiU

These are the documents provided by LiU:

- this document
- Appendix 1 TenderForm
- Appendix 2 MoU_HirlamB_license
- Appendix 3 AC2018/NWP18_summary
- Appendix 4 Definitions and Tables
- Appendix 5 A-40.1-10.pdf (Datacenter overview)
- Appendix 6 A-43.6-21.pdf (Datacenter overview)
- Appendix 7 E-61-1.pdf (Datacenter overview)
- Appendix 8 E-600.1-113.pdf (Datacenter overview)
- Appendix 9 E-600.1-116.pdf (Datacenter overview)
- Benchmark archive files

1.2.2 In the Tender

These are the minimum of documents provided by the Tenderer in the Tender:

- This document with all questions answered
- Appendix 1 TenderForm (see subsection 5.1.2)
- Appendix 3 AC2018/NWP18_summary
- Benchmark output, as presented in subsections 4.4.2, 4.4.3 and 4.4.4
- Performance and power consumption commitment report as presented in subsection 4.4.5.1
- Description of the tendered solution, subsection 1.1.7.5
- Certificates:
  - three (3) certificates from Tenderers located abroad (see subsection 5.1.3)
  - one (1) certificate from Swedish Tenderers (see subsection 5.1.4).

When the Tenderer refers to a document, LiU will be grateful if the reference is accompanied with a subsection identification or another identification, to simplify the reading.

2 Technical specification

2.1 Definitions

Where appropriate, e.g. referring to the capacity of computer memory, multipliers indicating powers of 2 are used. In all other contexts, e.g. storage capacity, multipliers indicating powers of 10 are used. Also see Appendix 4 Definitions and Tables.

2.2 General requirements

2.2.1 Continuous load

All components must be designed and manufactured for efficient continuous operation in a datacenter with an inlet temperature of 20-25 degrees Celsius and Relative humidity 20-80%

Fulfilled? (Yes/No answer)

2.2.2 Compatibility

All components must be compatible with CentOS-7 (www.centos.org) or RHEL-7 and its packages.

Fulfilled? (Yes/No answer)

2.2.3 Update process

If CentOS-7 compatible
CentOS-7, possibly with additional local patches and additions, will be installed and used on all servers. Delivered software components must have a documented update process compatible with an up-to-date CentOS-7 installation.

If RHEL-7 is required
RHEL-7, possibly with additional local patches and additions, will be installed and used on all servers. Delivered software components must have a documented update process compatible with an up-to-date RHEL-7 installation.

Which of the alternatives is chosen: CentOS-7 or RHEL-7? (Free text answer)

Is there a documented update process included? (Yes/No answer)

2.3 Servers

2.3.1 General

2.3.1.1 Servers A, B, and C There will be three different servers in the cluster:

Server A compute servers “small”

Server B compute servers “large”

Server C system servers - five (5) servers

FIXME: Account for GPU node config in AC20128 case.
FIXME: Specify for the NWP18 case system servers

2.3.1.2 Processor architecture Regardless of processor architecture, it must be based on 64 bits.

Fulfilled? (Yes/No answer)

FIXME: Discard or specify more closely

FIXME: If the processor architecture is not x86_64 some kind of small “convenience partition” is needed.

2.3.1.3 Connections All servers must be connected to the three (3) networks (Network A, Network B, and Network C) defined below.

Fulfilled? (Yes/No answer)

FIXME: Needs improvement

2.3.1.4 Maintenance port All servers must have a dedicated management port connected to Network C.

Fulfilled? (Yes/No answer)

2.3.1.5 Management All servers must have the ability to be managed via IPMI by any system server (Server C). This management must at least include System Event Log (SEL), Serial Over LAN (SOL), and Power Control.

Fulfilled? (Yes/No answer)

FIXME: A scriptable command line tool must be able to control the management processor wrt Power, SEL and SOL.

2.3.1.6 Net booting All servers must be configured for net booting (PXE or similar) via network B at delivery.

Fulfilled? (Yes/No answer)

FIXME: Specify which network(s) should be used.

2.3.1.7 If not PXE for booting If other method than PXE is used for booting, the solution must be described. Describe the solution (Free text answer)

2.3.1.8 Upgrading and configuring System ROM There must be an automatic and scalable method to upgrade and configure System ROM.

Fulfilled? (Yes/No answer)

FIXME: Incomplete, but left as is for now.

2.3.1.9 Airflow design All servers must have an airflow design such that outlet temperature is at least twelve (12) degrees Celsius higher than the inlet temperature when running Lifetime running cost verification, RELEVANT_COMPONENT component (see subsection 3.4.2.6).

Fulfilled? (Yes/No answer)

FIXME: This is valid for air cooled alternatives. Water cooled solutions will need different specification of the requirement.
FIXME: The Linpack reference will need to be updated. What should be used for "full load?"

FIXME: Water cooled solutions are possible but it has to take the entire cost of installation of a separate water cooling circuit up to a water–water heat-exchanger connected to the primary loop. Installations has to be in accordance with norms and standards and has to be cleared with Akademiska Hus, owner of the building. E.g. it has to include leakage detection and necessary water treatment.

2.3.1.10 System disk All servers must be equipped with an SSD (?) disk with a minimum storage capacity of at least 128 GB.

Fulfilled? (Yes/No answer)

FIXME: This needs further investigation!

2.3.2 Server A: Compute server “small”

Server A must be equipped with the following memory size:

- if the server has less than 16 processor cores per socket the server must be equipped with at least 24 GiB of primary memory per socket.
- if the server has 16 or more processor cores per socket the server must be equipped with at least 32 GiB of primary memory per socket.

Fulfilled? (Yes/No answer)

FIXME: Needs to be updated. A minimum requirement of a dual socket node is 48–64 GiB.

FIXME: Suggestion: Server A must be equipped with 48 GiB primary memory at a minimum. All available memory channels must be populated.

2.3.3 Server B: Compute server “large”

Servers B must be equipped with at least 256 GiB primary memory.

Fulfilled? (Yes/No answer)

2.3.4 Server C: System server

2.3.4.1 Processor and memory configuration All servers must have the same processor and memory configuration as Server B.

Fulfilled? (Yes/No answer)

2.3.4.2 Ethernet port for network C All servers must have one (1) Gigabit Ethernet port (1000Base-T) connected to Network C. This port is used by Server C to control devices connected to Network C.

Fulfilled? (Yes/No answer)

FIXME: Specify or clarify. Server C should have two connections to Network C. One to control the server and one to be used for the server to control other devices.

2.3.4.3 Unused Ethernet port All servers must have one (1) Gigabit Ethernet port (1000Base-T) unused.

Fulfilled? (Yes/No answer)

2.3.4.4 Ethernet port for network B All servers must be connected to Network B via one (1) ten-Gigabit Ethernet port.

Fulfilled? (Yes/No answer)

2.3.4.5 RAID-1 drives and disks All servers must have at least two (2) disk drives configured as one (1) RAID-1. Each disk drive must have a size of at least 200 GB (SAS or SATA) and must be hot swappable.

Fulfilled? (Yes/No answer)

2.3.4.6 Additional disk drive For each server an additional disk drive of the same type and size as in subsection 2.3.4.5 must be delivered as a spare.

Fulfilled? (Yes/No answer)

2.3.4.7 Power supplies All servers must have redundant and hot swappable power supplies.

Fulfilled? (Yes/No answer)

2.3.4.8 Serviceability All servers must be fully serviceable and replaceable without affecting the availability of any other server.

Fulfilled? (Yes/No answer)

2.3.4.9 PCI Express All servers must be configured to allow for later addition of one (1) midrange GPU using PCI Express 16x.

Fulfilled? (Yes/No answer)

2.3.5 Porting and testing servers

If the processor architecture is not based on x86_64 (AMD64 or Intel 64): - one server B and one server C for porting and testing purposes must be delivered three (3) weeks after the contract is signed. - the processors in these servers must be based on the same architecture as in the AC2018/NWP18 system, but can be of an older version.

Acceptable and fulfilled? (Yes/No answer)

2.4 Communication networks

2.4.1 General

2.4.1.1 General All servers must be interconnected using all the following networks:

- Network A: Fast application network
- Network B: General Ethernet network - Gigabit Ethernet
- Network C: Administrative network

Fulfilled? (Yes/No answer)

2.4.1.2 Redundant network equipment If a piece of network equipment (switches, etc.)

- is affecting more than 20% of all compute servers, or
- is connected to a Server C

it must be equipped with redundant and hot swappable power supply.

Fulfilled? (Yes/No answer)

2.4.2 Network A - Fast Application Network

2.4.2.1 Network A: general Network A will be used for communication between processes and threads within an application (e.g. MPI and TCP/IP), for communication with disk storage (e.g. Verbs and TCP/IP), and other system services.

2.4.2.2 Network A: TCP/IP gateway functionality Network A must include gateway functionality to bridge IPoIB on all servers to IPoEthernet externally.

Fulfilled? (Yes/No answer)

FIXME: This is a question for COS. The gateway may be unnecessary.

2.4.2.3 Network A: Unused ports There must be at least sixteen (16) unused ports in network A.

Fulfilled? (Yes/No answer)

FIXME: Contact Lenovo to ascertain whether they provide an OmniPath option for their GSS:es in our current setup. This will need to be included somehow in the offers.

2.4.2.4 Network A: Unused Gigabit Ethernet ports There must be at least eight (8) unused ten-Gigabit Ethernet ports with support for link aggregation (LACP) in network A.

Fulfilled? (Yes/No answer)

FIXME: Depends on the existance of a gateway as above

2.4.2.5 Network A: 10BASE-LR(LC) At least four (4) of the ports specified in subsection 2.4.2.4 must be equipped with 10GBASE-LR(LC).

Fulfilled? (Yes/No answer)

FIXME: Depends on the existance of a gateway as above

2.4.2.6 Network A: Gigabit capacity At least four (4) of the ports specified in subsection 2.4.2.4 must continuously be able to deliver at least five (5) Gigabit/s (unidirectional) from or to any one (1) server.

Fulfilled? (Yes/No answer)

FIXME: Depends on the existance of a gateway as above

2.4.2.7 Network A: Gigabit capacity, continued At least four (4) of the ports specified in subsection 2.4.2.4 must continuously be able to deliver at least thirty (30) Gigabit/s (unidirectional) from or to any eight (8) servers.

Fulfilled? (Yes/No answer)

FIXME: Depends on the existance of a gateway as above

2.4.3 Network B - General Ethernet Network

2.4.3.1 General Network B provides a generic Ethernet to all servers.

2.4.3.2 Network B: connections Network B must be connected to one (1) or more switches where a) the links between switches are at least ten (10) Gigabit/s b) any connection between two (2) servers pass no more than three (3) switches.

Fulfilled? (Yes/No answer)

2.4.3.3 Network B: Port level monitoring Network B must have port-level monitoring capability.

Fulfilled? (Yes/No answer)

2.4.3.4 Network B: VLAN capacity Network B must have VLAN capability.

Fulfilled? (Yes/No answer)

2.4.3.5 Network B: Connections to servers Connections to all servers must be one (1) Gigabit/s or faster. *Fulfilled? (Yes/No answer)*

2.4.3.6 Network B: connection to Server C Connections to all Server C must be ten (10) Gigabit/s or faster. *Fulfilled? (Yes/No answer)*

2.4.3.7 Network B: Unused Gigabit ports There must be at least twenty (20) unused ports capable of one (1) Gigabit/s or faster.

Fulfilled? (Yes/No answer)

2.4.3.8 Network B: Unused Gigabit ports, continued At least four (4) of the ports in subsection 2.4.3.7 must be capable of ten (10) Gigabit/s or faster.

Fulfilled? (Yes/No answer)

2.4.3.9 Network B: two ports 10GBASE-LR(LC) At least two (2) of the ports in subsection 2.4.3.8 must be equipped with 10BASE-LR(LC).

Fulfilled? (Yes/No answer)

2.4.4 Network C - Administrative Network

2.4.4.1 General Network C will be used to reach IPMI functionality and other management interfaces on all equipment from a separate network connection on Server C (see subsection 2.3.4.2).

2.4.4.2 Network C: external management Equipment with any form of external management via a network must be connected to Network C.

Fulfilled? (Yes/No answer)

FIXME: Ensure this is clarified wrt server C.

2.4.4.3 Network C: unused ports

There must be at least sixteen (16) unused ports.

Fulfilled? (Yes/No answer)

FIXME: The ports must be placed where servers C connect to network C. The same goes for network B. It is a matter of which network equipment is considered "central".

2.5 Racks and other additional equipment

2.5.1 Mounted in racks

All delivered equipment must be mounted in racks.

Fulfilled? (Yes/No answer)

2.5.2 Power distribution

Power distribution within the rack must be included. See description of the Datacenter in Appendix 5 - Appendix 9 for possible power connections.

Fulfilled? (Yes/No answer)

2.5.3 Balancing power distribution

Powered equipment must be distributed evenly over the three phases.

Fulfilled? (Yes/No answer)

2.5.4 Airflow in rack

Equipment must be mounted in rack for efficient airflow through the rack (front to back).

Fulfilled? (Yes/No answer)

2.5.5 Climate in rack

Equipment mounted from the back of the rack must be built to endure the demanding climate of the hot aisle.

Fulfilled? (Yes/No answer)

2.5.6 Sealing of racks

Racks must be carefully sealed to eliminate inefficient airflow through the rack.

Fulfilled? (Yes/No answer)

2.6 Installation

2.6.1 Power draw

At maximum load (see subsection 3.4.2.6 – Power usage measurements) any rack configured with servers type Server A or type Server B must have a power draw 15 kW or higher.

FIXME: Verify that 15 kW is appropriate.

Fulfilled? (Yes/No answer)

2.6.2 Installation of equipment

Hardware installation must be included. This includes e.g. unloading, unpacking, installation of all equipment in assigned space, cable management, labelling of cables, and recycling of used packaging material.

Fulfilled? (Yes/No answer)

2.6.3 Documentation

General

Documentation in English must be included in the delivery and regarded as part of the delivery.

The documentation must include

- specifications of components. These can be given as web-references
- specification of physical configuration: racks, servers, other components etc.

Information about the servers
Installation must include a sorted list in electronic format with necessary information to install and configure the system, including:

- MAC-addresses for network B
- MAC-addresses for network C including all managed equipment.

- password to service processor for all servers.

Fulfilled? (Yes/No answer)

2.6.3.1 Serial numbers Any serial numbers or identification information needed for future support processes, must be included in electronic form at installation.

Fulfilled? (Yes/No answer)

2.7 Service and support

2.7.1 Service level

Service on-site with Next Business Day (NBD) response time must be included for all equipment. For equipment or components where service on-site with NBD is not possible, there must be a corresponding alternative (e.g. spare parts on-site).

Fulfilled with at least one of the alternatives? (Yes/No answer)

In case of an alternative to NBD response time on-site, please describe the alternative solution (Free text answer)

FIXME: This requires clarification. Define importance tiers for different equipment and different service levels for them, and what we require for each instance. Define important equipment (servers C and connected equipment etc.). Define level of support we need for them.

FIXME: We should minimize the time NSC personnel spend on common service issues (such as replacing memory dimms, PSUs, etc.). I.e. we should spend minimum time on error reporting, fault diagnostics, spare part logistics, etc. for common errors. It is unclear whether Service personnel on-site or a cache of spare-parts on-site is the most efficient.

NEGOTIABLE

2.7.2 Troubleshooting

The Contractor is responsible for all advanced troubleshooting of equipment and components.

Fulfilled? (Yes/No answer)

2.7.3 Spare part logistics

The Contractor is responsible for spare part logistics.

Fulfilled? (Yes/No answer)

2.7.4 Availability of spare parts

The Contractor must guarantee availability to spare parts for the System for a period of at least six (6) years.

Fulfilled? (Yes/No answer)

2.7.5 All service and support costs are included

All costs for Service and support must be covered by the annual fee, including parts, work, travel, and living.

Fulfilled? (Yes/No answer)

2.7.6 Skilled personnel

The Contractor's personnel, servicing and supporting the System, must be highly acquainted with the hardware and software in question.

Fulfilled? (Yes/No answer)

2.8 Software

2.8.1 General

Intel Cluster Studio is fully licensed by LiU.
CentOS-7 is managed by LiU.
GPFS is managed by LiU.

2.8.2 Operating system

If RHEL-7 is required a RHEL-7 license for the system must be included in the tender.

Please answer "Yes" below if either

1. *RHEL-7 is not required*

2. *RHEL-7 is required and a RHEL-7 license is included in the Tender.*

Otherwise answer "No".

2.8.3 Other software

Compilers, MPI libraries, and other tools and libraries (except Intel Cluster Studio) necessary to acquire specified performance must be included in the Tender.

Fulfilled? (Yes/No answer)

2.8.4 Other system software/firmware

FIXME:

- identified
- compatible
- supported for system lifetime
- delivered
- working with a supported GPFS client.
- Work out GPFS connection to disk storage on:
  - OmniPath
  - Other stuff ... (drivers, monitoring etc.)

3 Delivery and installation

3.1 Directives

The complete system and any separate parts must comply with all appropriate EU directives, including CE marking, where appropriate. *Fulfilled? (Yes/No answer)*

3.2 Delivery dates

The Earliest installation date, i.e. when NSC will allow the Contractor access to the facilities, is EARLIEST_INSTALL_DATE. The date when the actual installation on-site will begin shall be set in the Delivery plan.

The System must be operational for the Purchaser and its customers at the latest LATEST_DELIVERY_DATE, i.e. Agreed delivery date.

Are these dates acceptable for the Tenderer? (Yes/No answer)

FIXME: Installation of AC2018 will be staged into two deliveries. The first early enough to take the installation into production July 1, 2018. This installation should be of a size that match the performance of existing system. THe second delivery can be installed after

1. all users from the existing system have been transfered to the first installation and
2. the existing system has been decommissioned and removed.

Installation of NWP18 will consists of two separate clusters installed in separate computer rooms. One of the clusters may be installed at SMHI in Norrköping.

3.3 Installation

Installation must be done in the Datacenter at NSC in Linköping, Sweden.

FIXME: For NWP18, this is only partially correct.

3.3.1 Information about the Datacenter

Additional plans of the installation site can be provided. Also see Appendix 5 to Appendix 9 for plans. Regarding site inspection see subsection 1.1.12.

Location
Linköping University
Campus Valla
Linköping

GPS coordinates: 58.3920, 15.5720

Construction
(Section plan of cell 1 and entrance: A-40.1-10)
The floor in the computer room (Cell 1) is made out of concrete. There is no installation floor, i.e. all connections (power, network, etc.) outside a single rack should be at the top of rack. Within the computer room there are two modules each with two rows of racks, encapsulating the two hot isles in the room.

(Detailed plan of module: A-43.6-21)
In each rack row there are two sections of 6150 mm of free rack space divided by a steel pillar of 80 mm. Maximum installation rack heigth is 2400 mm. Rack height of 2000 mm is preferred.

Power
(Cable trays: E-600.1-113, E-600.1-116; Cross-section of cell: E-6.1-1)

Power is distributed above racks on pre-installed trays at 3000 mm above floor. Power outlets are installed on the edge of the tray. Possible configurations: - 2 x 16A 3-phase outlets (IEC 60309) 230V/50Hz per rack with 16A fuse - 2 x 32A 3-phase outlets (IEC-60309) 230V/50Hz per rack with 25A fuse. Other power configurations are possible but will induce additional costs to the installation of the system.

All power in the computer room is connected via a group of five UPSes. The UPSes either run as double-converter or as hot standby (economy saving mode). Fans, pumps, and other supporting equipment are connected via a separate UPS. Attaching the equipment to power is done by LiU in collaboration with the Contractor.

Cooling
The computer room is built for air cooling only. Air for cooling equipment in rack is delivered in one end of the cold isles and is vented out at one end of the hot isles. Most air is recirculated. Installed racks with equipment must handle inlet air, into the front of rack, with a temperature of 20 - 25 degrees Celsius and a relative humidity (RH) of 20-80%. Unused airflow must be minimised. All installed equipment must be sealed against the racks. Airflow between and underneath the racks must be eliminated. Not properly sealed racks will require additional unused airflow and additional running costs in both fan power and chilled water.

FIXME: Water Cooling...

Fire protection
Sampling equipment for fire detection is used. For fire suppression a water mist system, Marioff HI-FOG, is installed.

Shipment and transportation
The computer room building is accessible by trucks with lengthy trailers. There is no loading dock. Necessary equipment for unloading heavy equipment must be provided. Maximum equipment cross-section for passing doors into the computer room: height: 2330 mm, width: 1400 mm.

Shipment on pallet with weights up to 2000 kg can be handled. Heavier equipment may be handled upon request and in agreement. Any additional costs for transportation are covered by the Contractor.

System wide connection
Internal communication between racks within a system should not use pre-installed trays in the building. Preferably, cable management is resolved using cable trays on top of rack and should be part of the installed solutions.

Noise levels
The noise level in the hot isles zone can be dangerously high for people. Ear

protection must be used in the hot aisles whenever computer systems in the isle are running. Ear protection equipment is available on-site.

Special requirements on the Datacenter
Any special requirements for installation at the Datacenter must be stated below. The cost for required modifications or additions are covered by the Contractor.

State any requirements (Free text answer)

3.4 Acceptance

3.4.1 General

Functional test is divided into six parts:

- Demonstration of T_tot,
- Runtime variability,
- Linpack stability test,
- STREAM,
- TCP/IP performance,
- Lifetime running cost verification.

3.4.2 Functional test

3.4.2.1 Demonstration of T_tot LiU will install the NSC software environment (see subsection 4.4.5) on the delivered AC2018/NWP18 system for the Functional test. Committed AC2018/NWP18 system performance T_tot is defined in subsection 4.4.1 and must be demonstrated during the Functional test. Requirements for benchmark runs must be adhered to, see section 4.

To achieve the pledged T_tot performance, the Contractor has the option of installing its own software environment.

If T_tot given in the Tender cannot be achieved during the Functional test, the Contractor must increase the capacity of the AC2018/NWP18 system at his own expense such that T_tot will be met.

Are these conditions acceptable by the Tenderer? (Yes/No answer)

3.4.2.2 Runtime variability Definitions:

- T_Arome_average: Arithmetic average of all Arome wall clock execution times from the Arome session used to calculate T_tot.

- T_Arome_single: Wall clock execution time for one Arome benchmark job, identical to one of the Arome benchmark jobs used for calculating T_tot, on an otherwise idle system.

Independent of job placement across servers, T_Arome_single must never exceed T_Arome_average by more than 7%.

Is the difference described above at most 7%? (Yes/No answer)

3.4.2.3 Linpack stability test The AC2018/NWP18 system must reach at least 75% of total theoretical peak performance by running a single copy of the Linpack HPL benchmark across the system. The run must include all cores on all delivered compute servers. Aggregated memory usage across all compute servers must be at least 70% of the total primary memory of the system.

The processor frequency must not go below nominal during the run. BIOS settings on the servers and other system settings must be identical to the BIOS and system settings used when benchmarking components for the T_tot calculations.

Instructions for running the Linpack stability test are found in the Linpack part of the benchmark package (Linpack_bench_system.v1.0).

Fulfilled? (Yes/No answer)

3.4.2.4 STREAM Memory bandwidth will be measured from the STREAM benchmark using at least 75% of system memory. No server shall run STREAM TRIAD at less than 90% of the median from all servers.

The benchmark can be downloaded from http://www.cs.virginia.edu/stream/

Fulfilled? (Yes/No answer)

3.4.2.5 TCP/IP-performance The following requirements must be fulfilled for the performance of the TCP/IP connection from Network A (Infiniband) to Ethernet.

Load is generated by running the iperf application (http://sourceforge.net/projects/iperf/) between pairs of compute and storage servers. Compute servers in AC2018/NWP18 and existing storage servers at NSC attached are to a HP-Procurve 5412 Ethernet switch. Performance measured by using port counters on the HP-switch. See figure in Appendix 4, Definitions and Tables.

Write performance from multiple servers Between eight (8) pairs of compute and storage servers a total performance of at least thirty (30) Gigabit/s from compute servers to storage servers must be accomplished. This is measured as an average over ten (10) minutes.

Read performance from multiple servers Between eight (8) pairs of compute and storage servers a total performance of at least thirty (30) Gigabit/s from storage servers to compute servers must be accomplished. This is measured as an average over ten (10) minutes.

Write performance from single server Between one (1) pair of compute and storage servers a performance of at least five (5) Gigabit/s from compute server to storage server must be accomplished. This is measured as an average over ten (10) minutes.

Read performance from single server Between one (1) pair of compute and storage servers a performance of at least five (5) Gigabit/s from storage server to compute server must be accomplished. This is measured as an average over ten (10) minutes.

Fulfilled? (Yes/No answer)

3.4.2.6 Lifetime running cost

Lifetime running cost for the system is the sum of lifetime energy cost and lifetime building rental cost. Lifetime energy cost is based on the estimated average energy usage for the system during four (4) years. Lifetime building rental cost is based on maximum possible energy usage for the system during four (4) years. The SEK/kWh price for energy and building rental are specified in Appendix 3.

Lifetime energy usage
Definitions:

- P_idle: Power usage on an idle AC2018/NWP18 system
- P_LP: Power usage on the AC2018/NWP18 system where all servers are running single server Linpack
- Load_factor=0.70. Typical load is estimated to be 0.70 of full load
- Average_application_power_use_vs_Linpack=0.83. An average application is estimated to use 0.83 of the power that Linpack is using.

P_tot = Estimated average power usage of the system
= P_LP*Load_factor*Average_app_power_use_vs_Linpack+(1-Load_factor)*P_idle
= P_LP*0.70*0.83+0.30*P_idle

Lifetime energy usage for the system is defined as the estimated average power usage times number of hours for 4 years: P_tot * 35040

Lifetime maximum possible energy usage
Lifetime maximum possible energy usage for the system is defined as power usage where all servers are running single server Linpack times number of hours for 4 years: P_LP * 35040

Power usage measurements
P_idle and P_LP must be measured:

- using an external power measurement equipment on the complete system
- as an average over 10 minutes
- during the following environmental conditions
  - Temperature: 24-25 degrees Celsius
  - Relative humidity: 20-60%
- using 240V AC, 50Hz
- with BIOS settings on the servers and other system settings identical to the BIOS and system settings used when benchmarking Arome and Nemo for T_tot calculations.

For P_LP, the following also applies:

- must be measured during the main compute part of the Linpack benchmark
- all compute servers in the system must run the single server Linpack test simultaneously
- each compute server must reach at least 75% of its total theoretical peak performance
- the run must include all cores on all delivered compute servers of using at least 75% of the primary memory on each compute server. (FIXME: Correct for any GPU servers offered?)

Instructions for running single server Linpack for power usage measurements are found in the Linpack part of the benchmark package "Linpack_bench_system.v1.0".

Lifetime running cost verification during Functional test
Lifetime running cost verification during the Functional test is done by measuring P_idle and P_LP according to the specifications above under the section "Power usage measurements".

The values of P_idle and P_LP are used in Appendix 3 to produce the Lifetime running cost. If Lifetime running cost is higher than presented in the Tender, the Contractor must cover the additional cost.

Are the conditions in "Power usage measurements" and "Lifetime running cost verification during Functional test" above fulfilled and accepted by the Tenderer? (Yes/No answer)

FIXME: Linpack is problematic to use as the reference for the power measurement as an unscrupulous vendor could in principle tune it to the worst performance to lower the power consumption. Is it possible to measure consistently during the benchmarking?

3.4.3 Stability test

3.4.3.1 General Stability test is divided into three parts:

- Error frequency
- Transient errors
- Availability.

LiU shall continuously and without delay inform the Contractor about errors occurring during the stability test period. The stability test period goes on until there is a continuous 60-day period where the System complies with the requirements.

3.4.3.2 Error frequency If a component ceases to function according to its specification it is counted as an error.

If a redundant component ceases to function according to its specification it is counted as an error.

In each component category, errors are counted separately.

Error thresholds
During the 60 day stability test period the number of errors must not exceed the thresholds in Appendix 4 Definitions and Tables, subsection 3.4.3.2.

3.4.3.3 Transient errors A server that is non functional after a restart but becomes functional after additional restarts is defined as exhibiting a restart error.

Ten (10) restarts of the complete system (Servers A,B, and C) must not generate more restart errors than one percent (1%) of the total number of restarts.

Example:
System has 90 servers, 10 system restarts = 900 restarts in total. No more than 9 restart errors are accepted.

3.4.3.4 Availability Availability testing is divided into the following three categories:

- System servers (Server C): Number of functional system servers connected to their three networks (Network A, B, and C)
- Compute servers (Server A, Server B): Number of functional compute servers connected to their three networks (Network A, B, and C)
- Network A and Network B (Communication network, Infiniband): Actual bandwidth that is delivered from all components in Network A and Network B expressed as a percentage of the full system bisection bandwidth.

A component is unavailable from the time that an error is reported until correct functionality can be verified on the component after an error correcting action has been performed. An error case number is received when the error report is made. A reported error that later is found not to be an error does not impact error frequency or availability.

Availability during the 60 day stability test period is measured as an instantaneous value as well as a time integral (time x availability) over the period. Availability must not go below the thresholds in Appendix 4 Definitions and Tables, subsection 3.4.3.4 Availability.

Example of instantaneous and time integral availability for compute servers: Assume a total of 90 compute servers in AC2018/NWP18. Test period is 60 days=1440 hours.

- Instantaneous availability of 85% for compute servers correspond to at least 77 compute servers available at any time during the test period.
- Time integral availability of 98% corresponds to at least 127 008 compute server hours out of 129 600 possible.

3.5 Schedule for delivery

When the Contract is signed the Agreed delivery date is set to DELIVERY_DATE.

When the Contractor has installed the System and run his own tests the System is handed over to the Purchaser to commence the Functional test. This test is estimated to run for 20 Working days and at its successful completion the Actual Date of Delivery is set, expected to be no later than Agreed delivery date.

Direct following Actual delivery date the Purchaser starts the Stability test, to be run until 60 days has been passed without violating thresholds.

FIXME: Either the Contractor or Purchaser can do the Functional test. For contractual reasons, it is easier if the Contractor demonstrate the actual performance of the installed system before handing over the System to the Purchaser at the Actual Date of Delivery.

4 Benchmarks

(DRAFT SPECIFIC NOTE! This section is written for AC2018 and is very much in a WIP state and will change upon vendor input and NSC internal revision. This section will look very different for NWP18. The gist of the section is summarised in the subsection 4.1)

4.1 Introduction

AC2018 may consist of possibly several parts; a bulk part for scientific data production, possibly equipped with GPUs to a significant extent, and a pre-/post processing, visualisation and development part. The pre-/post-processing part will need to be of a x86_64 architecture to support legacy software only available for this architecture. The bulk part of the system can be one of x86_64, POWER or ARM64/Aarch64 architectures.

AC2018 will be running a wide variety of applications in production. Three of these – VASP, GROMACS and CP2K – will be used for benchmarking of performance and will be included in the evaluation, see section 4.4. The benchmark applications are written in Fortran, C, C++ and CUDA and are parallel applications using MPI, or MPI and OpenMP for their parallelism.

For the bulk part of the resource, there is an option to offer CPU-only equipped nodes or a combination of CPU-only equipped nodes and GPU-equipped nodes. GPU-equipped nodes can be offered as part of the system to the amount of 30% of the budget (TCO). That is, a vendor can offer either a homogenous, CPU-only equipped cluster or a cluster with GPU-equipped up to a budget share of 30% and the remainder with CPU-only equipped nodes. Disregarding the GPU, GPU and CPU-only nodes are not required to be identically configured, e.g. CPU model (aside from its architecture) and socket count need not be the same for the different node types. Offer evaluation will be based on fastest time to solution on a job mix of the specified benchmarks, see section 4.4.

An important part of the benchmark process is that LiU want Tenderers to focus on configuring the hardware in an optimal way for the applications with respect to:

- processors/GPUs (architecure, type, clock, number of cores),
- memory (speed, bandwidth, number of ranks),
- interconnect (bandwidth, latency),
- power consumption.

Unless otherwise specified, all benchmarks run in the evaluation and verification phase must be run on machines with identical settings between individual runs,

i.e. the individual runs should not need some reconfiguration between runs with respect to BIOS or other firmware settings or kernel parameters.

4.2 General Information

The benchmark distribution comes with README files that describes the process of building and executing the benchmarks, along with any restrictions on what compile time optimisations are allowed. The applications are written in Fortran, C, C++ and CUDA and are parallel applications using MPI and MPI+OpenMP.

4.2.1 Access to benchmark package

The benchmark packages are access limited with user name and password, and they can be found at https://www.nsc.liu.se/about/procurement/AC2018. Access can be obtained by registering interest at http://www.tendsign.com/eng/, upon which the user name and password will be provided. Changes to the benchmark packages are announced to all registered vendors.

4.2.2 File I/O

File I/O in the benchmarks is done in main memory (/dev/shm). However, since non-local (i.e. not local server level) storage is not part of this procurement, for functionality testing a small global accessible file system is required for storing the benchmark input and output files, distribute data sets, and synchronise multiple job start up.

4.2.3 Versions

The benchmark archives for GROMACS and CP2K are self-contained, while the VASP benchmark archive does not contain the source distribution, since it cannot be redistributed by NSC. Obtaining this version is the responsibility of the Tenderer. The exact version of VASP to be used is specified in the README file distributed with the NSC benchmark package.

To keep track of changes LiU will provide a version number for each release of the benchmarks VASP, GROMACS and CP2K. If modifications are necessary to the benchmark package during the procurement, it will be communicated via TendSign to all suppliers without delay.

4.2.4 Allowed modifications to the benchmarks

Allowed benchmark modification are specified in the README file of each benchmark's archive.

4.3 Benchmark overview

The benchmark evaluation will only be carried out on the majority part of the system, i.e. all parts which are not counted as pre-/post processing parts. For reference in the CPU and GPU context, the large part of the cluster is divided into a CPU-only bulk part, called the bulk, constituting 70% of the TCO budget, and a complement part from the remaining 30% of the TCO budget.

The complement part is constituted by servers either equipped with GPUs as a vendor choice or CPU-only servers identical to those of the bulk part. The choice will be left to the Tenderer, and will ultimately depend on any performance over TCO benefits GPU equipped servers may bring over CPU-only servers. Servers hosting a GPU need not otherwise be identical to those of the bulk part, but must have the same architecture.

Only VASP and GROMACS will be benchmarked on the complement part and the benchmarks are identical on it irrespective of whether the nodes are GPU-equipped or not, allowing an apples-to-apples performance comparison between offers containing GPUs and those that do not. The bulk, CPU-only part only has one set of benchmarks – VASP (CPU benchmark), GROMACS (70% of the total number of GROMACS jobs) and CP2K.

4.3.1 VASP

- VASP: Vienna Ab initio Simulation Package, http://www.vasp.at/
- VASP version 5.4.1
- Different benchmarks for bulk and complement
- Hybrid density functional theory (HSE06) run on the bulk cluster part
- Standard density functional theory (PBE) run on the complement cluster part
- Running and building instructions in the README file of the benchmark package

4.3.2 GROMACS

- GROMACS: http://www.gromacs.org/
- GROMACS version 2016.3, optionally with architecture specific patches backported from gromacs git development branch
- Regular mdrun
- adh_dodec, protein in a dodecahedral box
- 95 000 Atoms
- Run on both CPU and GPU, 30 % of the jobs on the complement part of the cluster
- Minimum performance: 200 ns/24h trajectory simulated
- Running and building instructions in the README file of the benchmark package

4.3.3 CP2K

- CP2K: https://www.cp2k.org/
- CP2K version 4.1
- QS DFT *Ab Initio* Molecular Dynamics run
- 256 Water molecules
- Only run on CPU
- Maximum allowed wallclock run time: 120s
- Running and building instructions in the README file of the benchmark package

4.4 Evaluation

4.4.1 T_tot, committed AC2018 performance

The evaluation metric for the AC2018 procurement will be a composite figure measuring the total time, T_{tot}, to execute a mix of jobs from four benchmark cases; VASP – 2 benchmarks, GROMACS and CP2K. Lower T_{tot} means better (higher) performance.

The individual number of jobs for each benchmark type i, J_i, is predetermined and set by NSC. It may be assumed that the total run time for the VASP job mix will

be around twice those of GROMACS and CP2K (which have equal total job mix run time), giving VASP twice the weight in the evaluation compared to GROMACS or CP2K.

The total run time is calculated as:

$$T_{tot} = \sum_i T_i, \qquad i = \text{VASP-CPU, VASP-GPU, GROMACS, CP2K} \tag{1}$$

where T_i is the time taken for the offered cluster to run J_i jobs, where the subscript i denotes a specific benchmark i (VASP-CPU, VASP-GPU, GROMACS, CP2K).

Assuming long timescales, we can define job turnover rate r_i for an offered cluster, for a specific offer benchmark i as

$$r_i = \frac{N_o}{n_i \cdot t_i} \tag{2}$$

Here N_o is the size of the offered cluster part (bulk or complement) in number of nodes, n_i are the number of nodes needed to achieve wallclock run time t_i.

With total run time to run J_i jobs

$$T_i = \frac{J_i}{r_i}, \tag{3}$$

this gives us

$$\begin{aligned} T_{tot} &= \sum_i \frac{J_i}{r_i} = \sum_i \frac{J_i \cdot n_i \cdot t_i}{N_o} \\ &= \frac{1}{N_o} \sum_i J_i \cdot n_i \cdot t_i. \end{aligned} \tag{4}$$

This leaves for a vendor to find their optimum size of cluster (N_o), job width (n_i) for individual benchmarks and their respective wallclock run times (t_i, at least meeting the performance requirements). The n_i and t_i are solely influenced by the choice of servers and interconnect.

The acceptance testing of performance metrics n_i and t_i will focus on statistically verifying pledged figures on the delivered system fully loaded by benchmark jobs of all kinds, e.g. by means of sampling significant number of runs of each benchmark type.

4.4.2 Requirements for benchmark runs: VASP benchmark check list

Is a description of the system on which the benchmarks were run including processors, interconnect, memory size, and processor clock speed attached? (Yes/No answer)

Is a description of software and software versions, containing Linux version, compilers, MPI, and math libraries attached? (Yes/No answer)

Code changes are only allowed to the extent specified in the benchmark README file. If changes have been made accordingly, attach a summary of changes that was done. (Free text answer)
Present a summary of changes done to scripts, if any (Free text answer)
Are run scripts attached? (Yes/No answer)

Is build configuration file attached? (Yes/No answer)

Compute server sharing is not allowed. Does each job in the session have a unique set of servers? (Yes/No answer)

Huge pages must be disabled in the Linux kernel.

Fulfilled? (Yes/No answer)

The session should be run on a newly rebooted system.

Fulfilled? (Yes/No answer)

Are sessions run standalone? (Yes/No answer)

Are BIOS settings identical to BIOS settings used for the Linpack test for lifetime running cost verification? (Yes/No answer)

Are standard output from all runs attached? (Yes/No answer)

Are timing information files from all runs attached? (Yes/No answer)

4.4.3 Requirements for benchmark runs: GROMACS benchmark check list

Is a description of the system on which the benchmarks were run including processors, interconnect, memory size, and processor clock speed attached? (Yes/No answer)

Is a description of software and software versions, containing Linux version, compilers, MPI, and math libraries attached? (Yes/No answer)

Code changes are only allowed to the extent specified in the benchmark README file. If changes have been made accordingly, attach a summary of changes that was done. (Free text answer)

Present a summary of changes done to scripts, if any (Free text answer)
Are run scripts attached? (Yes/No answer)

Is build configuration file attached? (Yes/No answer)

Compute server sharing is not allowed. Does each job in the session have a unique set of servers? (Yes/No answer)

Huge pages must be disabled in the Linux kernel.

Fulfilled? (Yes/No answer)

The session should be run on a newly rebooted system.

Fulfilled? (Yes/No answer)

Are sessions run standalone? (Yes/No answer)

Are BIOS settings identical to BIOS settings used for the Linpack test for lifetime running cost verification? (Yes/No answer)

Are standard output from all runs attached? (Yes/No answer)

Are timing information files from all runs attached? (Yes/No answer)

4.4.4 Requirements for benchmark runs: CP2K benchmark check list

Is a description of the system on which the benchmarks were run including processors, interconnect, memory size, and processor clock speed attached? (Yes/No answer)

Is a description of software and software versions, containing Linux version, compilers, MPI, and math libraries attached? (Yes/No answer)

Code changes are only allowed to the extent specified in the benchmark README file. If changes have been made accordingly, attach a summary of changes that was done. (Free text answer)

Present a summary of changes done to scripts, if any (Free text answer)
Are run scripts attached? (Yes/No answer)

Is build configuration file attached? (Yes/No answer)

Compute server sharing is not allowed. Does each job in the session have a unique set of servers? (Yes/No answer)

Huge pages must be disabled in the Linux kernel.

Fulfilled? (Yes/No answer)

The session should be run on a newly rebooted system.

Fulfilled? (Yes/No answer)

Are sessions run standalone? (Yes/No answer)

Are BIOS settings identical to BIOS settings used for the Linpack test for lifetime running cost verification? (Yes/No answer)

Are standard output from all runs attached? (Yes/No answer)

Are timing information files from all runs attached? (Yes/No answer)

4.4.5 NSC software environment

The cluster will need to run the NSC software environment:

- The operating system is CentOS-7
- GPFS and Lustre file systems
- In-house developed software for installing, booting, monitoring, accounting, and management built on Python, Django, Nagios, BitTorrent, Dracut, collectl, and many other packages
- SLURM job management and job scheduling.
- Modules for managing software packages
- Compiler wrappers and script for compilation and execution of MPI-jobs

4.4.5.1 Requirements for benchmark runs: Performance estimates check list Details of analytical reasoning that underlies performance and power consumption estimates are included in a separate document "Performance and power consumption commitment".

Fulfilled? (Yes/No answer)

The report above should include a clear, concise, and reasonable description on how the commitments were reached.

Fulfilled? (Yes/No answer)

Does the report above include performance on existing system(s)? (Yes/No answer)

Does the report above include methods used to project performance from existing system to the offered AC2018/NWP18 system? (Yes/No answer)

Does the report above include a description on how the estimated and committed AC2018/NWP18 system performance T_tot (see subsection 4.4.1) and power

usage P_idle and P_LP (see subsection 3.4.2.6) were calculated for the offered system? (Yes/No answer)

Describe uncertainties in the calculations of the committed performance and power consumption. (Free text answer)

5 Qualification

5.1 Legal position

5.1.1 Legal position

According to the Swedish law of public procurement (LOU) the Tenderer may not:

- Be in bankruptcy or liquidation proceedings, under compulsory administration or subject to composition or have until further notice ceased making payments or be subject to trade prohibition,
- Be subject to filing for bankruptcy, compulsory liquidation, compulsory administration, composition, or other similar proceedings,
- Have been found guilty of unprofessional conduct according to final and conclusive judgement,
- Have been found guilty of serious wrongdoings in professional conduct,
- Have failed to fulfill liabilities concerning taxes and social fees.

Please confirm that none of the grounds for exclusion stated above is applicable. (Yes = confirmation). (Yes/No answer)

5.1.2 All Tenderers, Tender Form

Appendix 1 to this document, "Tender Form", must be signed and enclosed to the Tender, e.g. as a PDF.

5.1.3 Tenderers located abroad

Tenderers located abroad must enclose copies of the following national official documents to the Tender:

1. certificate of enrolment in a professional trade register,

2. certificate that the Tenderer has fulfilled obligations relating to the payment of social security contributions,
3. certificate that the Tenderer has fulfilled obligations relating to the payment of taxes.

5.1.4 Swedish Tenderers

Swedish Tenderers must enclose a copy of "Registreringsbevis" from Bolagsverket to the Tender. Linköping University will check Swedish Tenderers regarding payment of taxes and social security contributions towards Skatteverket.

5.1.5 Validity of information

None of the information requested above must be older than three (3) months from last day for submission of Tender.

5.2 Economic and Finacial Position

FIXME: Possible reqirement: revenue up to two times the value of the contract. ... liability insurance?

5.3 Technical ability and capacity

FIXME: Tenderer should have delivered a system listed on Top500 during the last three years.

Should this installation be in Europe? This may inhibit a few plausible Vendors. However, having the first large scale installation in Europe from the Vendor may cause difficulties and delays in the supply of spare-parts, support, and advanced service.

5.3.1 Reference

The Tender must include description of one (1) project, similar to this one, executed during the past three (3) years, with a contract sum higher than XX,000,000 SEK. This description must include reference specified with name(s), telephone number(s), e-mail address(es) to contact person or persons, and the contract sum. A Tenderer who lacks reference will not be considered.

Is at least one (1) reference included in the Tender? (Yes/No answer)

Give description of and contact details to reference(s) here: (Free text answer)

6 Suggested contract

FIXME: This Section has not been updated.

6.1 Parties and Intent of Contract

The following Contract has hereby as per agreement been signed between Linköping University, for National Supercomputer Centre, org. No 202100-3096 (below Purchaser) and [name] with org. No [no] (below Contractor). This Contract covers delivery of a HPC system in Linköping, including installation and Service and support.

6.2 Scope

This contract covers a HPC delivered and installed at NSC, Linköping University, in Linköping, Sweden, and a service agreement for the equipment covering the requested four (4) years, which can be extended one (1) plus one (1) year.

6.3 Definitions

See Appendix 4 Definitions and Tables.

6.4 Duration of the support contract

Service and support is included in the Contract for four (4) years, starting at Actual delivery day, which can be extended one (1) plus one (1) year.

6.5 Documents

This contract consists of the following documents and appendices. If any, colliding terms or interests according to this contract, the below order of priority is to apply:

1. Signed amendments to this Contract, if any
2. This Public Contract
3. Questions and Answers given in the Tender through TendSign.

4. Invitation to Tender (ITT) dated [datum]
5. The Tender from the Contractor, dated [datum]

Specifications, amendments, or exceptions in the Tender from the Contractor, which is not specified in the ITT documents, are not part of this agreement, unless there is a written amendment stating otherwise.

6.6 Contact persons

6.6.1 For Linköping University

Nnnnnn Nnnnnnn nnnnn.nnnnn@liu.se +nn nn nnnnnn

6.6.2 For [Tenderer]

Nnnnnn Nnnnnnn nnnn.nnnnn@aaaa.xxx +nn nn nnnnnn

6.7 Delivery

6.7.1 Delivery date and test periods

Earliest installation date is EARLIEST_INSTALL_DATE, as defined in Appendix 4 Definitions and Tables. Agreed delivery date is delivery. The day when the Functional test is successfully achieved is defined as the Actual delivery date. When Actual delivery date is passed the Purchaser begins the Stability test of the System while using it in daily operation for at least 60 days. Stability is reached when the System has met the requirements described in the ITT, subsection 3.4.2.1.

6.7.2 Delivery terms

The delivery terms are DAP on-site at Linköpings universitet in Linköping, Sweden, including freight, transport, insurance costs to site of installation, packing and removal of packing material, installation, setting the system into operation, testing, and documentation. (DAP = Delivered at place, according to INCOTERMS 2010).

6.7.3 Delivery address

Linköpings universitet Nationellt Superdatorcentrum House Galaxen, Datorhall Kärnhuset SE-581 83 Linköping Sweden

6.7.4 Delivery notification

The Purchaser must be kept informed about all stages of the delivery process. Especially the date when the physical delivery is planned must be signalled in reasonable time. Delivery shall take place during Office hours.

6.7.5 Delayed delivery

Delay in delivery is based upon the dates stated in subsection 6.7.1.

6.7.6 Notification of delayed delivery

If the Contractor finds that the agreed time of delivery cannot be met or appears unlikely to be met, the Contractor shall, without delay, notify the Purchaser of this in writing. Such notification shall include reason(s) for the delay and the expected date(s) by which delivery can be made. If delay in delivery is caused by the Purchaser, circumstances within the Purchasers control, or circumstances specified in Force Majeure subsection 6.15 and notification has been given as specified above, the time of delivery shall be extended to a time justified by the circumstances.

6.7.7 Delivery plan

As soon as the Contract is signed by both parties a Delivery plan is produced together. The Contact persons above are responsible for this action. The intension of the plan is to fulfil the delivery within required time schedule, as presented in the ITT. The Delivery plan shall be signed by both parties. Both parties must also establish the formal document(s) to be used to confirm important

milestones in the project, e.g. when Agreed delivery date is reached and Stability test is finished.

6.7.8 Access to datacenter

The Purchaser shall give the Contractor access to the installation site Office hours, unless otherwise agreed, at the latest at Earliest installation date.

6.8 Installation

6.8.1 Installation

The installation must include all necessary items and cover all associated expenses for the System. When the Purchaser starts the Functional test, planned to start 20 Working days before Agreed delivery date, the System shall be complete, as specified in the Tender.

6.8.2 Safety and health requirements

The system fulfills all legal regulations on safety and health requirements.

6.8.3 CE marking and conformity with EU regulations

The complete system and any separate parts comply with all appropriate EU directives, including CE marking where appropriate.

6.9 Acceptance: Functional and Stability tests

The Purchaser will verify the delivery by running Functional test, described in the ITT, subsection 3.4.2. When the Functional test is run successfully, the Purchaser immediately shall send a written confirmation to the Contractor's contact person. If the Functional test does not finish successfully, the Purchaser must without delay inform the Contractor, since a delay will probably affect The Actual delivery date. When the delivery has been verified, Stability test will be run to verify the quality of the system. The Stability test is described in the ITT, subsection 3.4.2.1. When the Stability test is run successfully, the Purchaser immediately shall send a written confirmation to the Contractor's contact person. The Purchaser must continuously and without delay inform the Contractor about errors occurring during the Stability test period. The Stability test period goes on until there is a continuous 60-day period where the System complies with the requirements in the ITT, subsection 3.4.2.1. In case the Functional test fail, the Purchaser shall, if feasible, still be entitled to use the System to the fullest extent possible. This will not be regarded as the Functional tests have been passed and will not affect the Actual delivery date.

6.10 Service and support

Requirements for Service and support are described in the ITT, subsection 2.7. If the Purchaser wishes to prolong the Service and support agreement it must be

done in writing at least six (6) months before valid Contract expires.

6.11 Subcontractors

If subcontractors are used in the delivery the Contractor has full responsibility for any and all work performed by subcontractors. The Contractor is also responsible to validate that all subcontractors can present the same documents as presented by the Contractor in the procurement:

For foreign subcontractors:

- certificate of enrolment in a professional trade register
- certificate that the subcontractor has fulfilled obligations relating to the payment of social security contributions,
- certificate that the subcontractor has fulfilled obligations relating to the payment of taxes.

For Swedish subcontractors:

- a copy of “Registreringsbevis” from Bolagsverket,
- and also verify the subcontractor regarding payment of taxes and social security contributions towards Skatteverket.

6.12 Actions for infringement

The Contractor is liable for any and all claims made against the Purchaser regarding infringement of third party Intellectual Property Rights associated with the use of the System, as well as any other costs arising as a result of such a claim.

6.13 Liability for faults and liquidated damages

6.13.1 Liability insurance

The Contractor is required to obtain a liability insurance concerning the goods and services under this procurement to cover possible costs caused by the Contractor or his subcontractors during installation and during Service and support agreement.

6.13.2 Delivery

If delivery, i.e. end of successful Functional test defined as Actual delivery date, has not been fulfilled within the agreed time (on Agreed delivery date) and the time for delivery has not been extended in accordance with subsection 6.7.6 last paragraph, the Purchaser is entitled to liquidate damages of one (1) % of the Contract sum, for each seven-day period that the delay lasts.

If delay in delivery only applies to minor items or positions, not effecting the overall capacity, the Purchaser's right to liquidate damages is limited to 0,1% of the Contract sum, for each commenced seven-day period that the delay lasts. The Purchaser's maximum entitlement to liquidated damages pursuant to the above is however 12% of the Contract sum, excluding VAT.

At own discretion, the Purchaser is entitled to credit accrued damages either by deducting damages for delayed delivery from the agreed price when making payment or by presenting the Contractor with a specific demand for payment. If such a demand is presented, payment shall be made within 30 days of the day the demand is raised.

If delivery is delayed, and solely related to reasons of which the Contractor has full responsibility, the Contractor must also pay interest on any sum paid in advance for a period corresponding to the delay. The interest rate shall equal the reference rate of the Bank of Sweden plus 8 percentage units.

If the delivery is delayed by reasons of which the Purchaser has full responsibility, the Contractor may present claim to the Purchaser for extended costs, depending on the delay.

6.13.3 Functional test

If the System does not reach the performance value, specified in the Tender, during the Functional test, the Contractor shall rectify the problem as described in the ITT, subsection 3.4.2.1. This rectification will require a new measure of power usage as described in the ITT, subsection 3.4.2.6. If the measure of power usage results in a higher value of Lifetime running cost than specified in the Tender, the Contractor shall cover the Purchaser's additional costs, as described in the ITT, subsection 3.4.2.6.

If the measure of power usage during Functional test results in a higher value of Lifetime running cost than specified in the Tender, the Contractor shall cover the Purchaser's additional costs, as described in the ITT, subsection 3.4.2.6.

6.13.4 Indirect loss or damage

A party is only liable for loss or damage that he has caused the other party through negligence or wilful act during the performance of the Contract. Such liability does not include compensation for indirect loss or damage, including but not limited to loss of earnings, loss due to the fact that the equipment cannot be used as intended, loss due to reduction, or loss of sales or production, loss as a consequence of damage to anything other than agreed delivery or similar loss.

6.14 Payment

6.14.1 Public contract price

The Contract sum is: [as stated in the ITT] SEK excluding Swedish VAT. The total cost for the system for four years is divided into four components (below). The tree first adds up to the Contract sum and the fourth is Lifetime running cost:

- hardware and installation: X,XXX,XXX SEK
- service and support for four years: X,XXX,XXX SEK
- other costs for the System: X,XXX,XXX SEK
- lifetime running cost: X,XXX,XXX SEK

6.14.2 Payment terms

The System shall remain the property of the Contractor until fully paid.

All payments are to be against invoice 30 days net. Invoicing and service charges will not be accepted. Interest on overdue payment, if any, will be payable in accordance with Swedish law.

Payment terms are as follows:

A. Following Actual delivery day 80% of System cost + Other costs ("System cost" and "Other costs" in Appendix 3) and 25% of Service and support cost ("Service and support cost for 4 years" in Appendix 3) are entitled to be invoiced by the Contractor.

B. Following completed Stability test 20% of System cost + Other costs are entitled to be invoiced by the Contractor.

C. Once a year from year two 25% of the total Service and support cost are entitled to be invoiced by the Contractor 12, 24 and 36 months after Actual delivery date.

6.14.3 Invoice address

Invoices will specify Linköpings universitet, D No. LIU_DIARY_NUMBER, and the name of the goods (or services) to which the delivery relates.

Invoices are sent to:

Linköpings universitet
Fakturasupport
Ref: 2310 [name] FIXME
SE-581 83 Linköping
Sweden

6.15 Force majeure

A party shall be released from his obligations if he proves that an impediment outside his control occurred, which he could not reasonably have foreseen or anticipated at the time of the sale and whose consequences he could neither have reasonably avoided nor overcome.

If the delay is due to someone engaged by a party to complete or partially perform the sale, such party shall only be released from liability to pay damages if the person whom he engaged would have been released from liability according to the first paragraph. The same applies if the delay results from a supplier engaged by the seller or anyone at an earlier stage in the chain of sale.

For a party to be entitled to claim a ground for release in accordance with the above, he shall without delay notify the other party of the occurrence thereof, and similarly of its cessation.

A party shall notify the other party of when it is estimated that performance can be affected.

In the event of force majeure on the part of the Purchaser, he shall compensate the Contractor for the additional costs that the Contractor may be caused in order to safeguard and protect the goods.

6.16 Cancellation

This contract is valid from the moment it has been signed by both parties. Either of the parties shall have the right to give notice of the immediate cancellation of the Contract if one party violates the stipulations in this Contract or does not follow the existing applicable laws, ordinances and/or stipulations and if the breach of contract is not merely one of minor consequence and also if the breach

has not been rectified within thirty (30) days after a written request has been made for this to be done.

The Purchaser is entitled to cancel all or part of the contract with immediate effect

a) if delivery is not performed within ninety (90) days from the Agreed delivery date, and the delay is not due to the Purchaser or circumstances on their part, or

b) if the System does not fulfil the Stability test requirements within one hundred and eighty (180) days from the Actual delivery date, or

c) if delivery is defective and this is not due to the Purchaser or circumstances on their part and the defect may be regarded as a substantial breach of contract, or

d) if the Contractor fails to perform other conditions of the contract and this may be regarded as constituting a substantial breach of contract, or

e) if the Contractor goes into liquidation or receivership, has entered into composition with its creditors, until further notice has suspended its payments or has an injunction against carrying on a business, or

f) if the Contractor is subject to a filing for bankruptcy, liquidation or receivership, or other such proceeding, or

g) if the Contractor has been convicted, and the conviction has acquired legal force, for a crime regarding his exercise of profession, or

h) if the Contractor is guilty of serious misconduct regarding his exercise of profession and the contracting entity can prove that, or

i) if the Contractor has not fulfilled its obligations regarding social security fees or taxes.

If cancellation is made, all payments made on the delivered platform will be refunded to the Purchaser within 30 days and the Purchaser shall have the right to claim damages.

If cancellation is made by the Purchaser without the right to cancellation according to the Contract, the Contractor shall have the right to claim damages.

6.17 Updates and changes

Changes to this contract can only be made through written amendments, signed by both parties.

6.18 Applicable law

Swedish law shall govern the Contract and disputes between the parties shall be determined by a Swedish court.

6.19 Marketing

The Contractor may not use the Purchaser as a named reference for marketing purposes without written acceptance from the Purchaser according to the “Act on names and pictures in advertising” (1978:800).

6.20 Signatures

_____

Linköping 2014-mm-dd [Location] 2014-mm-dd

_____ Signature

_____ Name, Title

_____ Signature

_____ Name, Title

6.21 Suggested Contract

This suggested Contract is presumed to form the basis for an agreement between the parties. Objections and additions from the Tenderer will be considered by LiU and possibly be included.

Present objections and additions to the suggested Contract (Free text answer)