Document number	WP2.005.065.020-TD-003
Revision	0.2
Author	Yogesh Wadadekar et al (see below)
Date	2011-10-24
Status	Submitted

MONITORING AND CONTROL DESIGN CONCEPTS SUMMARY

Name	Designation	Affiliation	Date	Signature
Submitted by:				
Yashwant Gupta	Project lead	NCRA	2011-10-24	
Yogesh Wadadekar, Yashwant Gupta, Niruj Mohan Ramanujam, Jayaram Chengalur, Jitendra Kodilkar, R. Balasubramaniam, N. Swaminathan, Amrit Lal Ahuja, Tejaswini Jana, Piyush Laad, Sunu Engineer, Vivek Mohile				
Accepted by:				
T.J. Stevenson	System Engineer	SPDO	2011-10-24	TJS
Approved by:				
Peter Dewdney	Project Engineer	SPDO	2011-10-24	TBR

DOCUMENT HISTORY

Revision	Date Of Issue	Engineering Change Number	Comments
0.1	2011-10-07	-	Draft released for external review
0.2	2011-10-24	-	Final version submitted to CoDR panel

DOCUMENT SOFTWARE

	Package	Version	Filename
Wordprocessor	MsWord	Word 2010	05_WP2-005.065.020-TD-003v0.2_dcs
Block diagrams			
Other			

ORGANISATION DETAILS

Name	National Centre for Radio Astrophysics
Physical/Postal Address	National Centre for Radio Astrophysics Tata Institute of Fundamental Research Post Bag No: 3 Pune University Campus, Pune India 411 007
Fax.	+91 20 25692149
Website	www.ncra.tifr.res.in

TABLE OF CONTENTS

Contents

LIST OF ABBREVIATIONS

AA..... .... Aperture Array
AD...................... Applicable Document
ALMA................. Atacama Large Millimetre Array
API..................... Application Programming Interface
ASKAP............... Australian Square Kilometre Array Pathfinder
ASTRON............ Netherlands Foundation for Research in Astronomy
CoDR................. Concept Design Review
EPICS................ Experimental Physics & Industrial Control System
EVLA .... Expanded Very Large Array
LOFAR............... Low Frequency Array for Radio-astronomy
LRU .... Line Replacement Unit
M&C................... Monitoring and Control
Mbps.................. Mega bits per second
PLC.................... Programmable Logic Controller
QoS .... Quality of Service
RFI..................... Radio Frequency Interference
SCADA .... Supervisory Control And Data Acquisition
SKA .... Square Kilometre Array
SKA1 .... SKA Phase 1
SKA 2 .... SKA Phase 2
SPDO .... SKA Program Development Office
TBC .... To Be Confirmed
TBD .... To Be Determined
TBF.................... To Be Filled In
TBW................... To Be Written

GLOSSARY

Central M&C: The portion of the M&C system that is located at the Operations & Maintenance Centre, and responsible for management of the system as a whole rather than a portion of it.

Common Operating State: An abstraction of the actual state of Components so that different Components can be treated uniformly by Regional M&C from the viewpoint of determining validity of transitions and validity of particular operations in a given state. The Common Operating State captures the health and usage status of entities.

Domain M&C: An auxiliary M&C system that manages resources belonging to a particular domain e.g. signal transport equipment, power equipment. This is developed (or acquired) by that domain and is outside the primary M&C hierarchy.

Entity: A generic term that may span any granularity from Element to Component to Part.

Local M&C: Controller for a Component typically developed by the Component provider, possibly off-the-shelf.

Monitoring points: Data items (including logs and reports), alarms or events that are forwarded up the system periodically or on-occurrence from the point of origin. Monitoring points may be processed, combined or abstracted by receivers to generate new monitoring points.

Operating Mode: A property of entities that reflect who can control the entity, and its relationships to other entities. In *Integrated* mode, the entity is fully functional and under the control of its parent nodes / operators. In *Offline* mode, the entity is under the control of maintenance engineers and not available to participate in normal operations. In *Commissioning* mode, the entity is not yet operational, has a lesser level of trust, and operates such that it does not interfere with the functioning of other entities (e.g. logical isolation of Commissioning network). In *Testing* mode the entity is partially functional and is available to participate only in test operations. The *Testing* and *Commissioning* modes affect the handling and propagation of alarms from the entity.

Regional M&C: M&C functionality that has primary responsibility for management of a portion of the system. This could be a station, a portion of the Core, or the collection of sensors and actuators in the system that is not associated with any station or Core.

Standardised Component Interface: All Components that interface with Regional M&C must conform to a standardised interface definition created by M&C. This includes not only interaction protocols at the software, hardware and technology levels, but also governance constraints on functionality and behaviour.

Sub-array: A logical partitioning of array of receptors into sub-groups to carry out different observations in parallel and control them concurrently.

System M&C: The parts of M&C developed by the M&C team i.e. Central M&C and Regional M&C.

1 Introduction

1.1 Scope of the document

This document relates to the SKA Monitoring and Control Domain Element and its Sub-elements. It is of a maturity commensurate with a Concept level of definition of the M&C Domain and the SKA Observatory as a whole. It should be noted that the Monitoring & Control design is aimed at the full SKA system and not only SKA1, since it is difficult to incorporate major architectural changes into a system after it is built, and because most of the eventual capabilities will be needed for SKA1.

1.2 Purpose of the document

The purpose of this document is to identify the major design issues in the design of the SKA Monitoring & Control Element and its Sub-elements, identify candidate solutions, and synthesise them into architectural alternatives, with an evaluation of their relative strengths and limitations. The objective is to analyze the feasibility of meeting requirements and determine the further work needed to firm up requirements and develop the SKA M&C architecture.

2 References

2.1 Applicable documents

The following documents are applicable to the extent stated herein. In the event of conflict between the contents of the applicable documents and this document, **the applicable documents** shall take precedence.

[1] SKA Phase 1 System Requirements Specification, T. Stevenson *et al*, SKA Project Document - WP2-005.030.000-SRS-002.
[2] SKA Science Working Group, "*The Square Kilometre Array Design Reference Mission: SKA Phase 1*", report, v.1.3, January 2011.
[3] K. Cloete et al, 'Strategies and Philosophies', document WP2-005.010.030-TR-001, Rev F.
[4] 'SKA1: High Level System Description', P.E. Dewdney et al, WP2-005.030.010-TD-002, 2011-02-14, Rev A

2.2 Reference documents

The following documents are referenced in this document. In the event of conflict between the contents of the referenced documents and this document, **this document** shall take precedence.

[5] Monitoring & Control Strategy, document WP2-005.065.000-R-001
[6] M&C Requirements for each Domain: Information Template WP2-005.065.050-RFI-001
[7] Inputs to M&C Development WP2-005.065.050-RFI-002
[8] Monitoring and Control Element Level Requirements, WP2-005.065.020-SRS-001.
[9] Monitoring and Control Design Concepts Description, WP2-005.065.020-TD-002.
[10] Monitoring and Control High-level Description, WP2-005.065.020-TD-001.
[11] 'Architecture for Secure SCADA and Distributed Control System Networks', Juniper Networks, Inc., 2010, white paper
[12] 'PVSS II', http://lhcb-online.web.cern.ch/lhcb-online/ecs/PVSSIntro.htm
[13] 'Overview of the Experimental Physics and Industrial Control System: EPICS', S.A. Lewis, LBNL, April 2000, A note for prospective users.

3 Comparative Evaluation of Candidate Architectures

This section briefly summarises the two candidate architectures identified, provides a comparative evaluation of their suitability for SKA M&C, and identifies a preferred architecture along with a way forward.

3.1 Summary of Candidate Architectures

Two alternative candidate architectures have been identified for M&C: Hierarchical Semi-Autonomous Control and Service Capability Matching. These architectures differ primarily in the way the M&C tiers (Central M&C / Regional M&C, Regional M&C / Local M&C) interface with each other.

Hierarchical Semi-Autonomous Control is the traditional approach and current state-of-the-art in control systems architecture. The tiers form a strict hierarchy: the Central M&C node directs the functioning of Regional M&C nodes, which in turn direct the functioning of Local M&C nodes (which themselves direct the functioning of their subsystems and so on). Moreover, this hierarchy is pre-defined at design time or at system deployment. Each parent node has pre-programmed knowledge of its child nodes, or at least of the types of child nodes that it can control. It has pre-programmed intelligence to handle the specific set of alarms generated by each child node, and to pass them up to its parent node if necessary. This concept has proven to be scalable, robust and can be augmented slightly to eliminate single points of failure and allow plug-n-play of child nodes.

The second architectural concept is more futuristic. It has a service-oriented flavour, with each node being viewed as providing a set of services to its parent node, utilising services from its child nodes. Nodes compose themselves dynamically into a hierarchy using service discovery and commitment concepts. Each node is aware of the capabilities and QoS it can provide, and can update these if it has internal failures. Parent nodes are viewed as service consumers, who can describe the services they need with the desired QoS parameters, and accept commitments from any node that can provide the desired services. This architecture has the potential to be more flexible and robust, and support more intelligent fault handling. In particular, since nodes know the effect of failures on QoS, and they know the target (desired) QoS, they can determine when the targets cannot be met, and can provide metadata augmentations that reflect the precise gaps between target QoS and actual QoS. This can be of considerable assistance in science interpretation.

3.2 Comparative Evaluation

The Hierarchical Semi-autonomous Control architecture is proven, scalable and widely used in radio astronomy. Its limitations are not serious:

- The single point of failure limitation can be overcome by providing failover for parent nodes in the hierarchy. Fortunately, the number of processors needed to handle the region load makes it relatively economical to provide failover capabilities even at the region level. Temporary power and communications backup can be provided to allow a safe shutdown of the region if the infrastructure fails.
- Plug-n-play capabilities are not important to SKA. The commissioning of new receptors or even adding new infrastructure equipment is a sufficiently major and effort-consuming activity that the need to configure Regional M&C with details of the new Component is not a significant burden. Further, it is possible to support limited plug-n-play (new Components, but not new types of Components) with hierarchical semi-autonomous control. Non-dynamic parent-child relationships are also not a serious limitation: configuration can be used effectively to cater for the flexibility required by the SKA in an Hierarchical Control architecture.
- More intelligent fault management is not a pressing need. Making fault handling and observation scheduling decisions on simple rules related to high-level abstractions ("at least 95% of the receptors are available") appears to be sufficient to meet system and science goals. The marginal value of more sophisticated decision-making is probably offset by the error-prone nature of more complex rules.

For these reasons, hierarchical semi-autonomous control is the conservative, low risk and preferred architectural alternative.

It is not known whether the Service Capability Matching architecture is even feasible. It is much more complex, QoS composition is known to be challenging to accomplish successfully for any domain, and not demonstrated for machine control. As such, it is not realistically available as a choice today. However, it has significant potential benefits that are hard to ignore. The ability to do superior metadata augmentation, perform more intelligent fault and resource management, determine feasibility of concurrent commitments and improve resilience to failures would make it a rather attractive option if it were feasible.

The most important consideration is that it is aligned with the direction of technology evolution. Given the widespread adoption of service-oriented architectures (for example, MeerKat is using a service-oriented architecture for its M&C), the general interest in QoS composition, and the fact that implementation of SKA M&C will not begin for several years, it is worth keeping an eye on the evolution of the technology landscape. Further, it turns out that the differences between the two architectures do not have a great impact on the rest of the M&C system design. This makes it feasible to contemplate evolving from one architecture to the other, if the latter were to become feasible at a still later date. It would entail significant changes to M&C software in each tier, but it is doable.

The recommendation therefore is to adopt the hierarchical semi-autonomy architecture, but to monitor the evolution of the technologies involved in Service Capability Matching. Hybrid architectures, for example incorporating service-orientation but not QoS composition, are also a possibility. The final decision on this will be made during the preliminary design phase.

4 Design Framework Summary

The candidate architectures are only a part of the M&C design concepts. Many of the core concepts for the M&C design are largely invariant with respect to the candidate architectures, which primarily affect the way M&C tiers interact with each other.

M&C design concepts are discussed in [10] and [9]. This is a brief overview of the proposed design framework:

- Layered approach, consisting of higher-level functions (Observatory Applications) running on a platform (Observatory Operating System).
- Three tiers: Central M&C, Regional M&C, Local M&C.
- Principle of semi-autonomy: Hierarchy of parent and child nodes, from the overall System Controller down to Local M&C of individual entities. Parent nodes provide goals (commands, set points), child nodes decide how to achieve the goals.
- Standardised interfaces to the Local M&C of all Components, defined common responsibilities and principles.
- Operations support online databases and operational integrations.
- Architectural principles for addressing engineering concerns.
- Information abstraction based on the concept of common operating states, health statuses and operating modes, with drill-down support. Common concept of action/reaction on alarms and events as far as possible, so that there is a uniform approach to situation handling throughout M&C. Incorporate this into the Standardised Component Interface, possibly providing standard libraries to ensure further consistency and avoid duplication.
- Remote management consoles for the Local M&C hierarchy and Regional M&Cs.
- Soft real-time control with synchronised time based execution between Regional and Local M&Cs.
- Ethernet as field bus for M&C as far down (towards hardware) as possible.

5 Architectural Variation Points

In addition to the choice among the candidate architectures, there are several other architectural variation points with respect to which choices need to be made during the preliminary design stage:

- M&C data storage at regional M&Cs: Each Regional M&C may produce several hundred thousand monitoring data points, including detailed internal state and event information for all of its Components, higher-level information derived from this raw data, logs and test reports. This may amount to about 50,000 monitoring points and around 50 Mbps data per region [9] based on an order-of-magnitude estimate, in the absence of hard information about numbers of Components in each region and monitoring points per Component at this stage of SKA design.

  The choice is whether to propagate all this data to Central M&C or to store it locally. The advantage of propagation to Central M&C is that it is possible to take advantage of shared infrastructure for storage and archival e.g. separate archive access bypassing Central M&C, easy physical access to the archive storage devices etc. The advantage of storing and archiving data at the regions is (a) reduced network traffic between the regions and Central M&C (b) significant reduction in computational capacity needed at Central M&C to handle the additional data points, leading to cost savings. Likely improvements in archive performance, since servicing all requests from Central M&C is a common bottleneck. (c) Possible simplification of design from the software platforms perspective: it is unclear if the software platforms can scale to the millions of data items that will be needed at Central M&C if all the data were to be made available there. There is relatively little difference from a usage perspective – drill-down for operators, maintenance engineers doing troubleshooting and other users – it will be transparent to them that some of the data retrieval requests go out to regions while others are serviced from Central M&C. On the other hand considerably more computing power and storage will be needed at Regional M&C, and there is also the operational problem of maintaining the additional hardware in remote locations. The security of data stored at remote locations may also be lower, leading to higher probabilities of data loss.

  Based on the above tradeoffs, storing data locally at Regional M&Cs appears to be an attractive alternative. Further exploration is needed to evaluate the tradeoffs in more detail and make a choice; at present, local storage in Regional M&Cs is the preferred alternative, with only summary data being transferred to Central M&C.
- Localisation of safety aspects and implications for backup networking: The current design concept proposes that all responsibility for safety monitoring and handling reside locally with each region. This proposal rests on the assumption that all safety threats are local. It also assumes that fail-safe design of individual Components, in addition to region-level monitoring and handling of safety threats, will be sufficient to eliminate any critical single-point-of-failure issues with respect to safety (i.e. no single failure will be sufficient to create an unsafe situation). It is unclear if this is entirely true – for example, there may be region-level safety threats (including threats to equipment) that are not specific to individual Components, such as severe weather or intruders. FMECA will be needed to establish whether in fact it is true that all safety threats are localised, and whether there are any threats that suffer from single-point-of-failure problems. Even if single-points-of-failure do exist, they can potentially be handled by adding an independent safety monitor system in each region, complementary to the M&C system. This, however, is an expensive option, since it will need a completely independent set of sensors, networking and power infrastructure to avoid common failure modes.

If it is indeed possible to localise all handling of safety threats, it confers multiple potential benefits: (a) Central M&C does not need to be treated as a safety-critical system. (b) Continuous connectivity between Central M&C and regions becomes less critical from a safety point of view. (c) It increases region autonomy, simplifying the design. In particular, it is sufficient to have an on-demand low capacity backup communication link between Central M&C and regions, rather than a fully fault-tolerant network architecture.

- Use of SCADA systems: There is a choice on whether to SCADA [11] packages based on distributed control databases such as PVSS [12] and EPICS [13] or merely to use general-purpose object-oriented languages such as C++ and Java. There are several advantages to SCADA packages: (a) they are widely used and proven for M&C implementation, and they are designed specifically for this purpose (b) they have excellent scalability (c) development effort is much lower with these packages. The advantages of general-purpose languages include (a) wide user base, leading to greater availability of libraries and off-the-shelf software (b) wider base of available expertise. The latter advantages are not very pronounced, because there is considerable off-the-shelf software available for the SCADA packages that exactly matches the domain needs, and they also have significant user bases and a moderately wide base of available expertise. On balance, SCADA packages supplemented with object-oriented technologies for UIs, algorithms and interfacing functions are currently the preferred option, with the choice of specific platform being subject to evaluation. It should be noted that these are not really alternative choices: it is likely that the solution will contain both technologies – the primary choice is whether the entire solution is developed manually using object-oriented technologies, or the base control framework is a SCADA system.
- Standardisation of Local M&C software platform: Local M&C systems will usually be developed by Component providers: suppliers of receptors, cooling equipment, power reticulation equipment, local power generation equipment etc. In some cases, the equipment will be off-the-shelf, and come bundled with M&C. In other cases, they may be customised or custom-developed for SKA. In that case, the M&C for these Components will be either developed afresh or modified from M&C solutions developed by the providers for previous projects. M&C already proposes to standardise the interface for Local M&C of Components; this variation point is about the possibility of standardising the internal implementation technologies for Local M&C as well.

  There are two alternative approaches to standardisation of the software platform to be used to implement Local M&C. The first possibility is to completely standardise the Local M&C software platform and require all Component developers to use it to build their M&C [except of course that the M&C software for any fully off-the-shelf Component will not be modified – instead, adapters will be required to be provided that map their interface to the Standardised Component Interface]. This might be quite challenging and expensive for the providers (costs that will eventually be borne by SKA), since they may have to completely re-engineer their M&C software. The other approach is to strongly encourage providers to use the standard software platform, but permit deviations in cases where this would cause them undue hardship. This creates a risk: if within the SKA lifetime a Component provider goes out of business, the responsibility of maintaining the M&C software for that Component type rests with SKA. Over time, SKA could find itself maintaining software [and hardware] that spans a wide range of technologies.

  The preferred approach is the latter one of providing flexibility to the providers to make choices, but strongly encouraging them to use the standard platform. This can be done both contractually, and by making available a significant volume of reusable components and guidance that make it convenient for them to use the standard platform. The lifetime technology risks would appear to be minor compared to the upfront costs, the possibility of dissuading or losing vendors who cannot afford major redevelopment of software, and the possibility of lower reliability with redeveloped software than with software that has been proven on their Component over multiple previous deployments. However, the alternative viewpoint that SKA is big enough to insist on technology commonality also has considerable support, and this aspect will need further investigation to make a choice one way or the other.

6 M&C Cost Estimation

This section attempts a rudimentary scoping of cost sources for the fully developed and deployed M&C system. At this stage, it is very difficult to attempt an actual cost estimate, however the ingredients of cost can be identified.

The costs of M&C realisation come from several sources:

1. Detailed design and realization of Regional M&C and Central M&C software, including integration interfaces.
2. Support to Local M&C and Domain M&C development teams.
3. Efforts for M&C integration and commissioning.
4. Cost of M&C computational infrastructure (processors, storage).
5. Cost of development of Local M&C systems.
6. Hardware costs for sensors/actuators, fieldbuses and data acquisition boards.
7. Cost of non-Component sensing equipment (weather sensors, security monitoring equipment etc).
8. Regional safety monitor, if an independent safety monitor for the region is considered desirable.

Of these costs, the first three items would be included within the M&C budget. The fourth item, computing infrastructure for M&C should probably be included in the M&C budget, though actual platform decisions and procurement are part of the Software & Computing domain. Development of Local M&C systems and associated sensor and hardware infrastructure will be part of the procurement of Components of which the M&C system would be a part, so its cost is included within other domains. It is unclear whether the cost of other sensing equipment would be part of the facilities domain. If a Regional safety monitor is deemed necessary, that too would be part of M&C development costs.

The infrastructure costs to be included within the M&C budget include processors and storage, additional engineering efforts and materials needed for installation and deployment of M&C. The processing infrastructure costs depend on the number of monitoring points, which is not known at this time. If a node has N monitoring points, and a single processor is capable of handling M monitoring points, then the number of processors needed is (N / M + 2 + k), where the 2 represents processors for control and the storage interface, and k represents a redundancy factor (typically 2 or more). With current technology, M is of the order of ~2000 for PLCs and high-end processors [depending on complexity of processing and several other characteristics], but this can be expected to increase over time. The storage needs to be large enough to hold at least a month of engineering data, and possibly more. It is too early in the design to identify the materials needed for installation and deployment of M&C, given that much of M&C installation and deployment will be part of Components deployment.

There will also be infrastructure costs for all the M&C processing other than monitoring: present GUIs, provide access to archived data, do observation planning, observatory management, proposal preparation, provide database servers, run schedulers, etc. Processing infrastructure costs will also include lab systems for development, simulations and integration. M&C infrastructure will include storage costs, possibly both at Central M&C and Regional M&Cs. Operator stations and display systems should also be included, unless they are already included in facilities costs.

This leads to the following preliminary breakdown of the cost sources for M&C development:

- People effort for
    - Engineering the M&C solution, including the definition of all its interfaces, including the interfaces to other systems and defining the Standardised Component interface.
    - Development of M&C software. Testing, installation and deployment of the M&C solution.
    - Guidance and support to Local M&C development, including possibility certification.
    - Support for commissioning and deployment of other systems, since M&C will be central to all commissioning and integration efforts.
- Infrastructure costs. This will consist primarily of

  - Computing infrastructure in each region, including processors and storage. A preliminary order-of-magnitude estimate would be ~10 processors and ~1PB of storage.
  - Associated networking, sensor and other infrastructure.
  - Computing infrastructure at Central M&C. This could be up to an order of magnitude more than at each region, depending on some of the choices and tradeoffs.
  - Sensor costs for any sensors that will be the responsibility of M&C.
- The regional safety monitor could be potentially relatively expensive, since it would be a safety-critical item. The cost would be not merely for the hardware, software and all the costs to verify and certify it, but also all the redundant sensors and associated wiring to independently connect them to the sensor.
- Costs for software licenses (including modelling and development tools), hardware setups for integration and testing.

6.1 M&C development for SKA1

The design of M&C is aimed at SKA2, since it will be difficult to retrofit scalability considerations into the architecture, and most M&C features are needed for SKA1.

The following is a representative list of M&C features whose implementation can be deferred to SKA2, though it is preferable to design for them in SKA1:

- Implementation of configuration and fault management functionality for Component types that are not part of SKA1.
- Storage at Regional M&Cs. For SKA1, it is likely that there is plenty of spare communications bandwidth. It may be simpler to eliminate the problem of deploying and maintaining archives in each region by pushing all the data to Central M&C, though it would have costs in terms of additional processing power.
- Independent safety monitors is another feature that can be considered for deferment. The lesser scale of SKA1 could mean fewer safety concerns related to remote outlying stations having safety problems that must be dealt with fully automatically. Emergency travel to SKA1 outlying stations may be relatively quick and first-level safety mechanisms may suffice, rather than requiring ultra-reliable and ultra-safe automated operation. This option requires an explicit project decision.
- Dealing with software platform architectural issues related to scale: partitioning of the data into multiple namespaces with brokers to create a transparent subscription space for users.
- Alarm filtering for intrusive vs. non-intrusive presentation. The number of alarms in SKA1 may be sufficiently tractable so that possibly all significant alarms can be presented to the operator.

7 Conclusion

A set of concepts that address the key driving requirements of SKA M&C have been presented in [9]. There are a couple of candidate architectures, one which is traditional and conservative, and another which is futuristic and may possibly become feasible within the SKA timeframes. There are also several architectural variation points that need to be resolved in the preliminary design phase.