The First Workshop on HPC Power Management: Measuring Effectiveness

From Facility to Component Level - A High Performance Computing Power Application Programming Interface

22 Sept. 2015

James H. Laros III
Principal Member of Technical Staff
Sandia National Laboratories
http://powerapi.sandia.gov

Sandia National Laboratories

Exceptional service in the national interest

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND2015-8115 C

Sandia National Laboratories

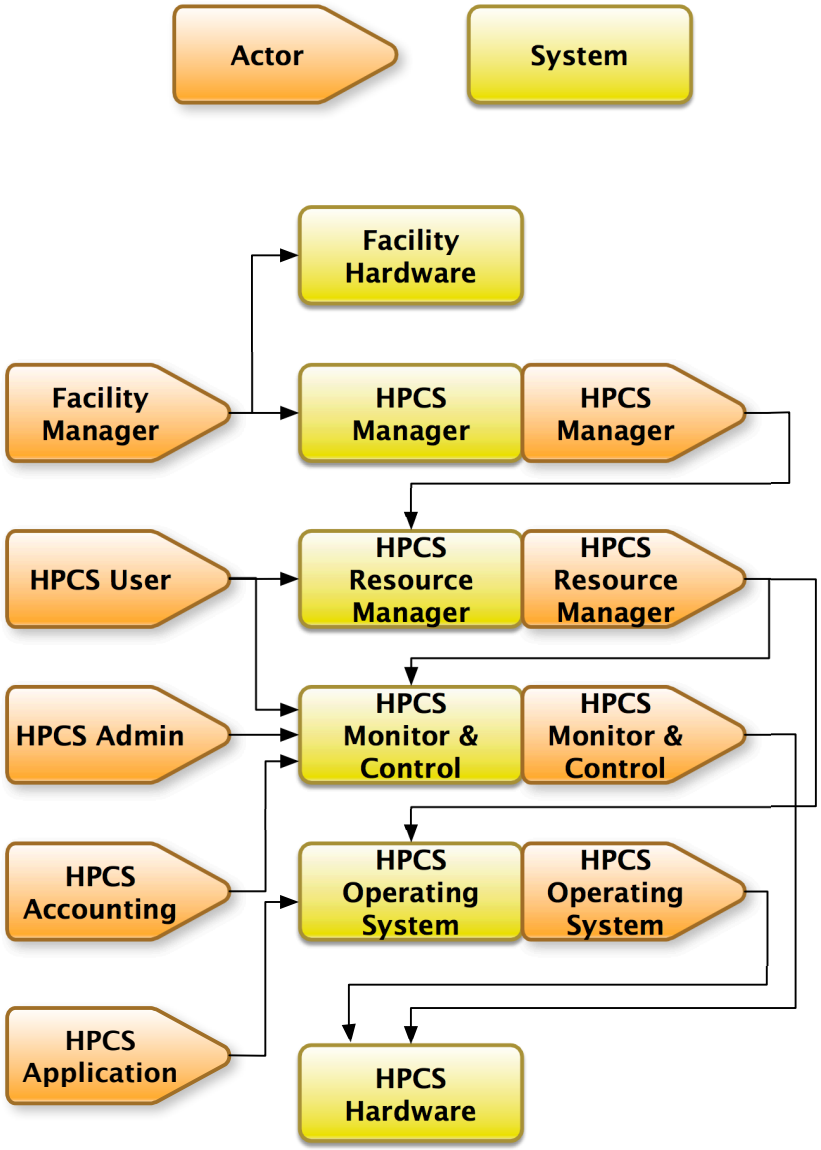
Early Work (2004)

Next Steps

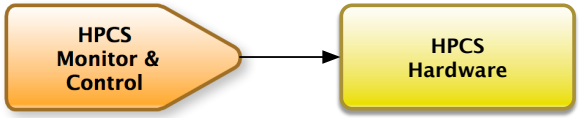
A Use Case Approach

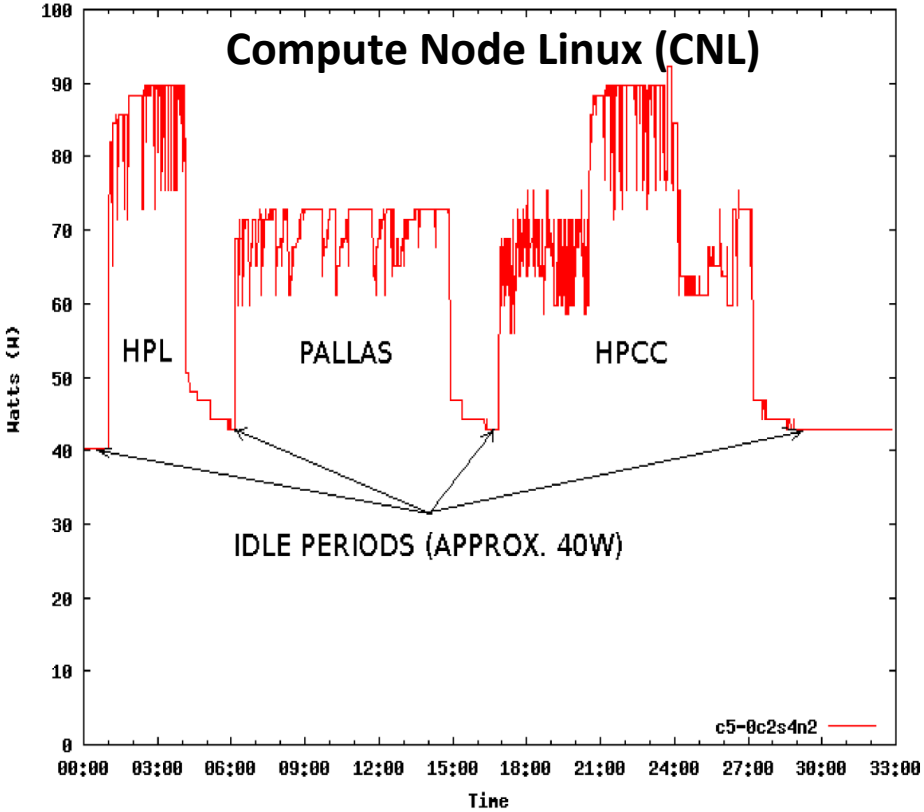
HPC Power API Specification

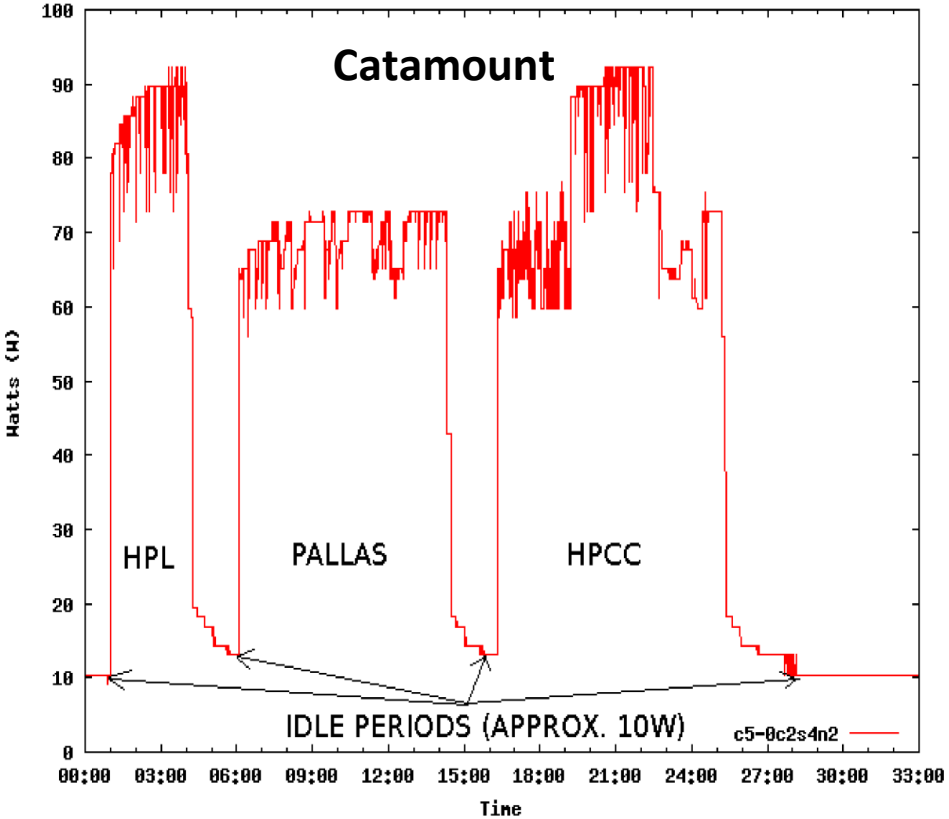
Sneak Peak

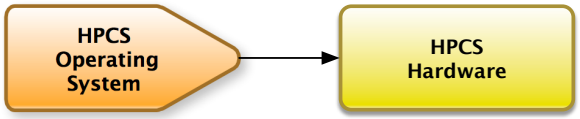

Actor
System
Facility Hardware
Facility Manager
HPCS Manager
HPCS Manager
HPCS User
HPCS Resource Manager
HPCS Resource Manager
HPCS Admin
HPCS Monitor & Control
HPCS Monitor & Control
HPCS Accounting
HPCS Operating System
HPCS Operating System
HPCS Application
HPCS Hardware

Application Power Profiles



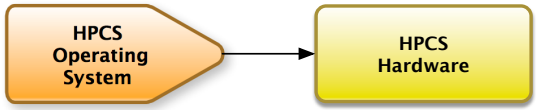


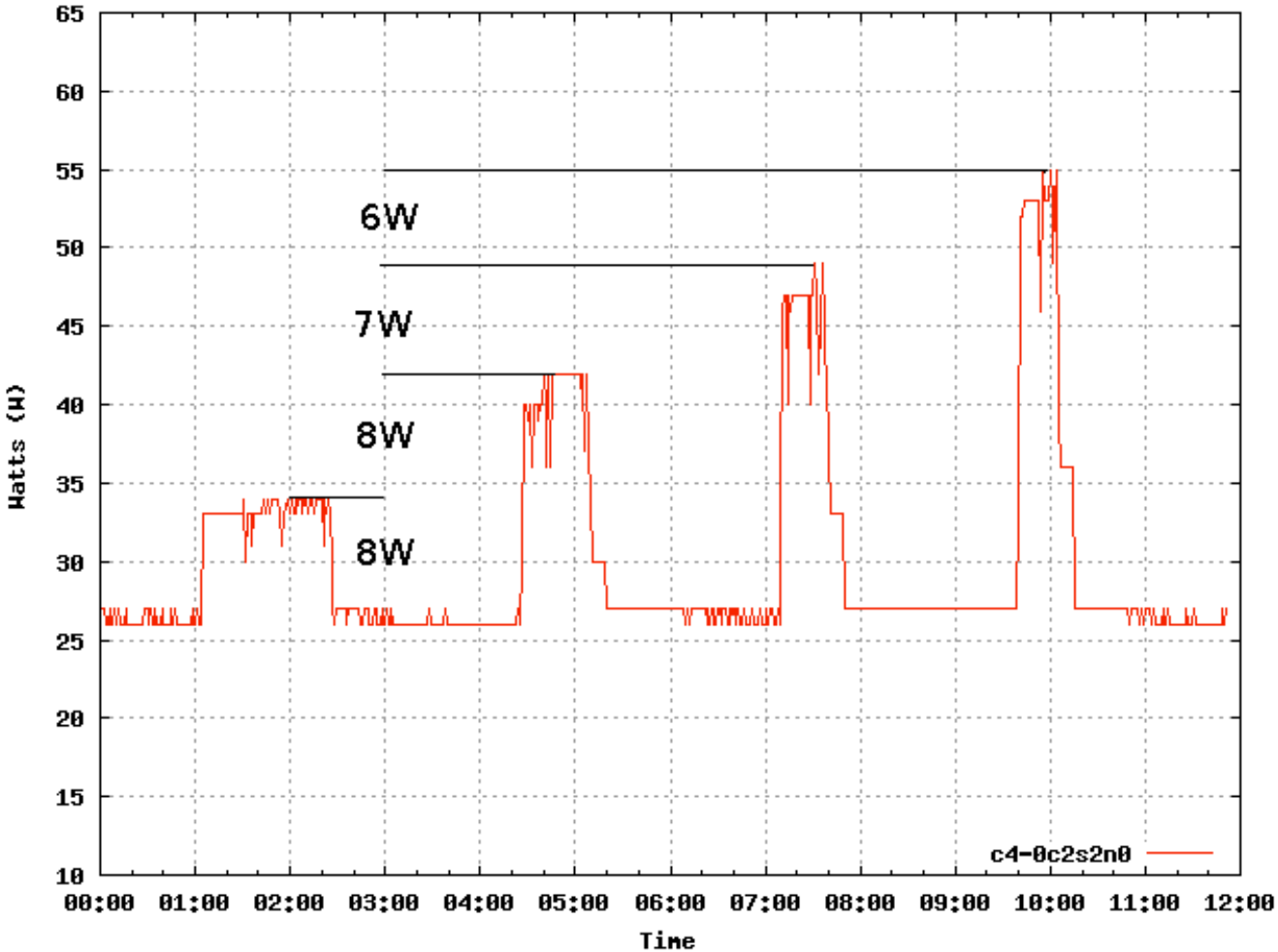


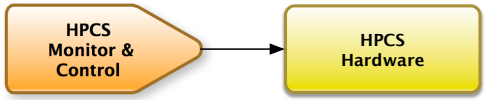


Idle Power

Deterministic Control





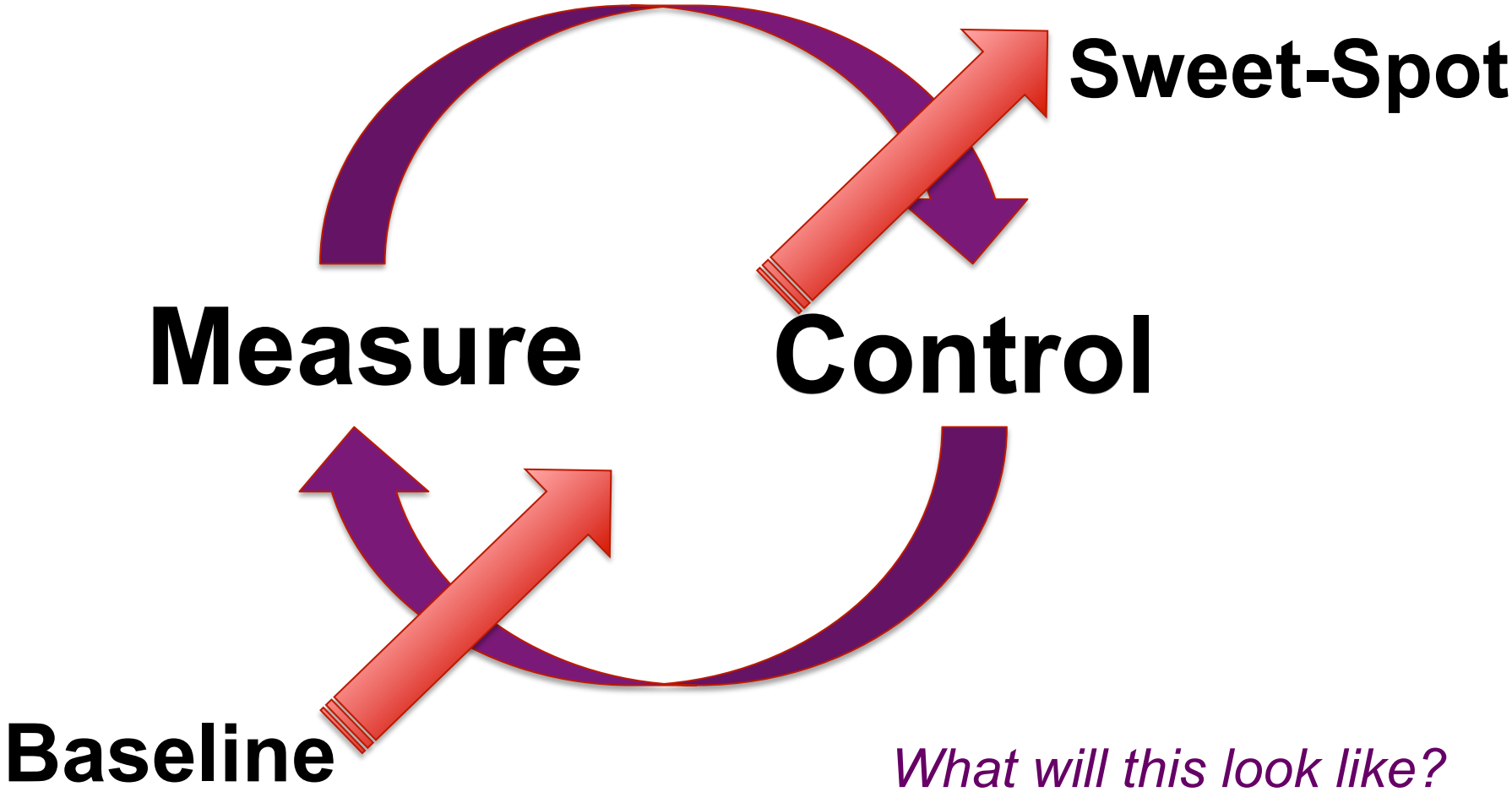


Ability to Observe

Results of CPU Frequency Tuning

	Nodes/Cores	P2 Run-time %Diff	P2 Energy %Diff	P2 Run-time %Diff	P3 Energy %Diff	P4 Run-time %Diff	P4 Energy %Diff
HPL	6000/24000	↑ 21.1	↓ 26.4				
Pallas	1024/1024	↑ 2.30	↓ 43.6				
AMG2006	1536/6144	↑ 7.47	↓ 32.0	↑ 18.4	↓ 57.1	↑ 39.1	↓ 78.0
LAMMPS	4096/16384	↑ 16.3	↓ 22.9	↑ 36.0	↓ 48.4	↑ 69.8	↓ 72.2
SAGE	4096/16384	↑ 0.402	↓ 39.5				
SAGE	1024/4096	↑ 3.86	↓ 38.9	↑ 7.72	↓ 49.9		
CTH	4096/16384	↑ 14.4	↓ 28.2	↑ 29.0	↓ 38.9		
xNOBEL	1536/6144	↑ 6.09	↓ 35.5	↑ 11.8	↓ 50.3		
UMT	4096/16384	↑ 18.0	↓ 26.5				
Charon	1024/4096	↑ 19.1	↓ 27.8				

What Did We Learn?



What will this look like?

Early Work (2004)

Next Steps

A Use Case Approach

HPC Power API Specification

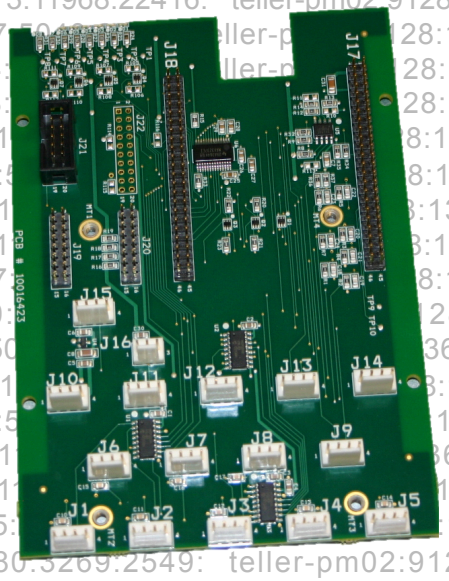
How do we proceed?

- We need platforms that have power and energy measurement capability
  - Red Storm (XT platforms) unique at the time -- we need ubiquitous
  - Need finer grained – want discrete CPU, Memory etc.
- Influence high end through RFP requirements
  - EEHPC Working group
  - Doesn't give us a way to continue our research **NOW**
  - Platform specific ok but limits architectures we can investigate
    - not ubiquitous
- Develop a commodity power measurement device
  - Works with "any" architecture
  - Will allow us to prototype our developing API theories
  - Enables progress **NOW**

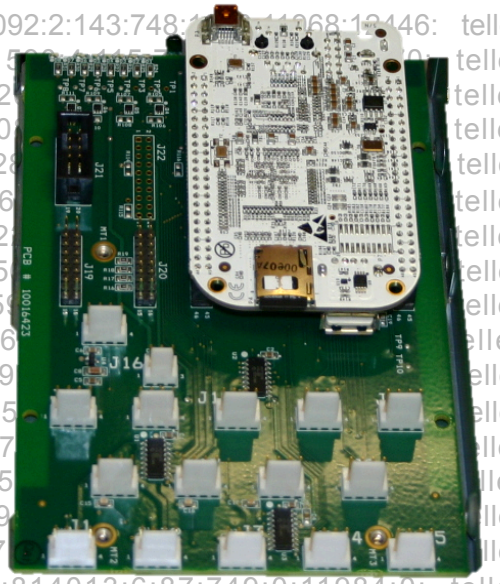
Sandia/Penguin Collaboration

small example of co-design

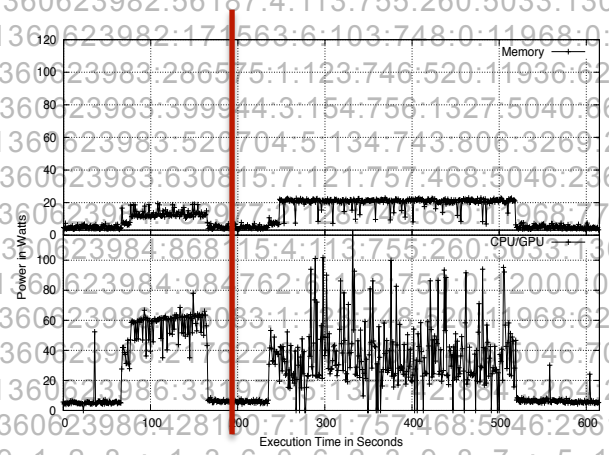
Commodity

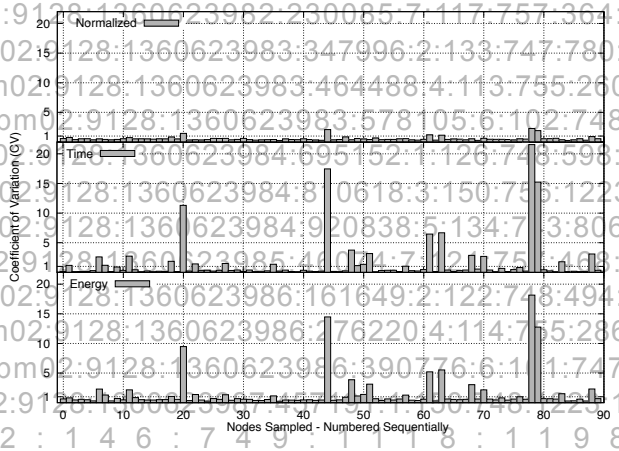


Custom PCB



PowerInsight





Early Work (2004)

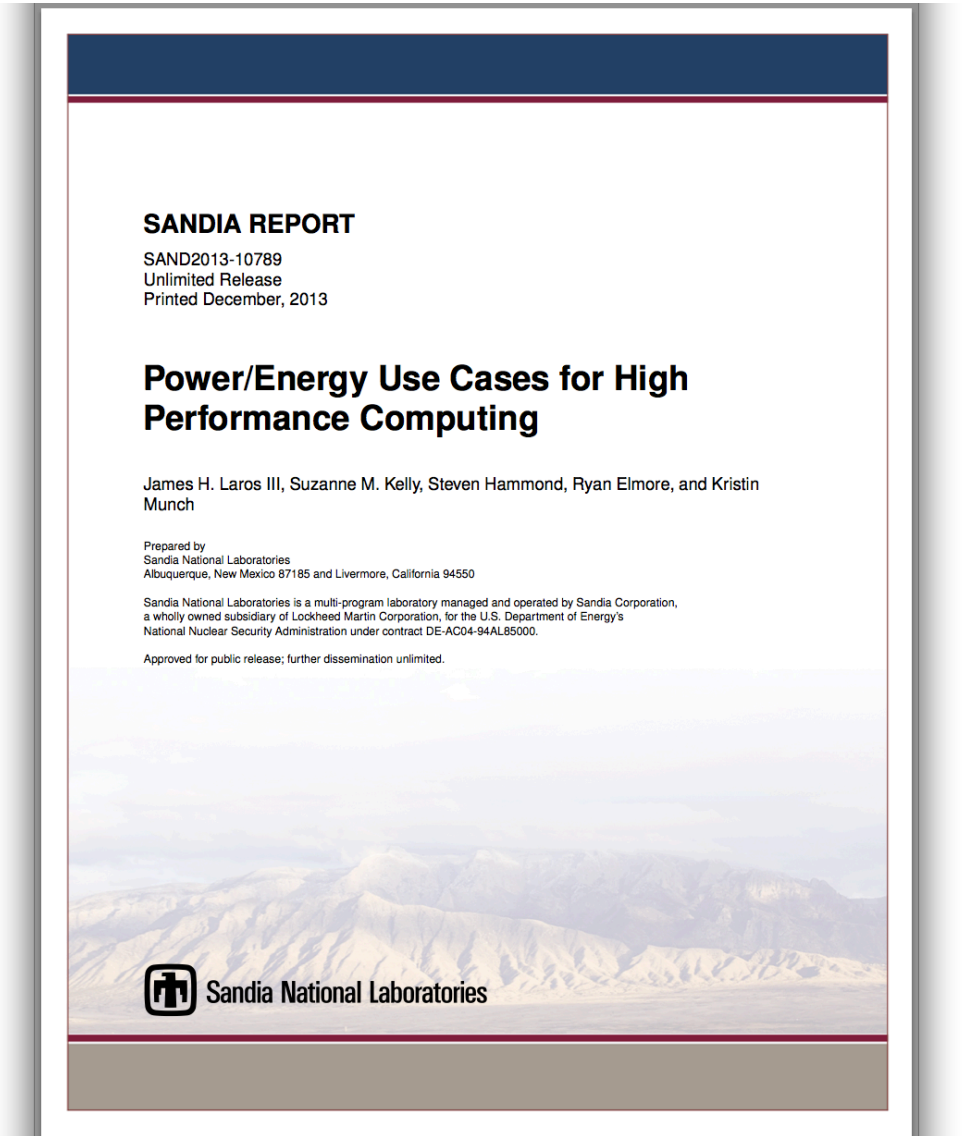
Next Steps

A Use Case Approach

HPC Power API Specification

Preliminary Investigation

- Sandia and National Renewable Energy Laboratory (NREL)
  - Sandia – Platform focused
  - NREL – Facility focused
- Primary Objectives
  - Identify SCOPE of Power API
  - Identify Roles
  - Identify Interfaces
- Product:
  - *"Power/Energy Use Cases for High Performance Computing"*
  - Use Case Approach
  - **Actor** (Roles) -> **System** pairs
  - Interaction between **Actor** and **System** defines an Interface
  - Facility to Component (hardware) **Scope**



SANDIA REPORT

SAND2013-10789
Unlimited Release
Printed December, 2013

Power/Energy Use Cases for High Performance Computing

James H. Laros III, Suzanne M. Kelly, Steven Hammond, Ryan Elmore, and Kristin Munch

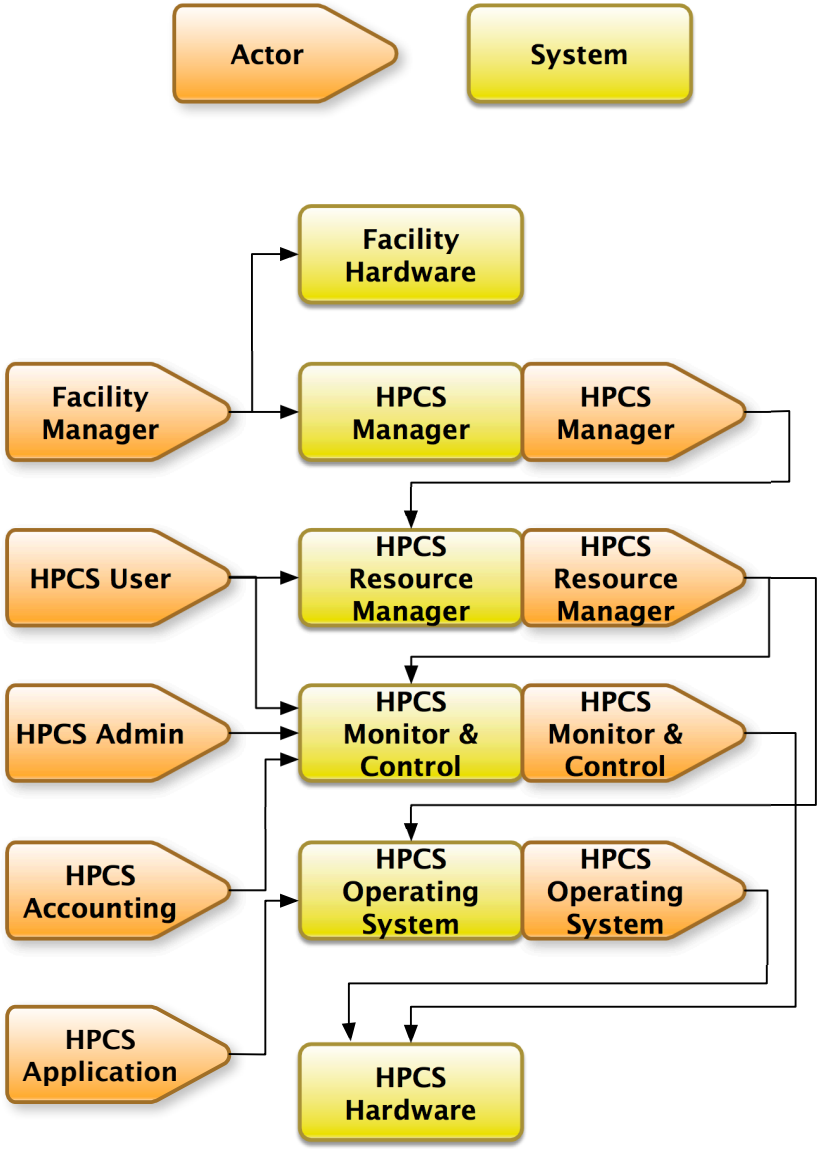
Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Result


Actor
System
Facility Hardware
Facility Manager
HPCS Manager
HPCS Manager
HPCS User
HPCS Resource Manager
HPCS Resource Manager
HPCS Admin
HPCS Monitor & Control
HPCS Monitor & Control
HPCS Accounting
HPCS Operating System
HPCS Operating System
HPCS Application
HPCS Hardware

Early Work (2004)

Next Steps

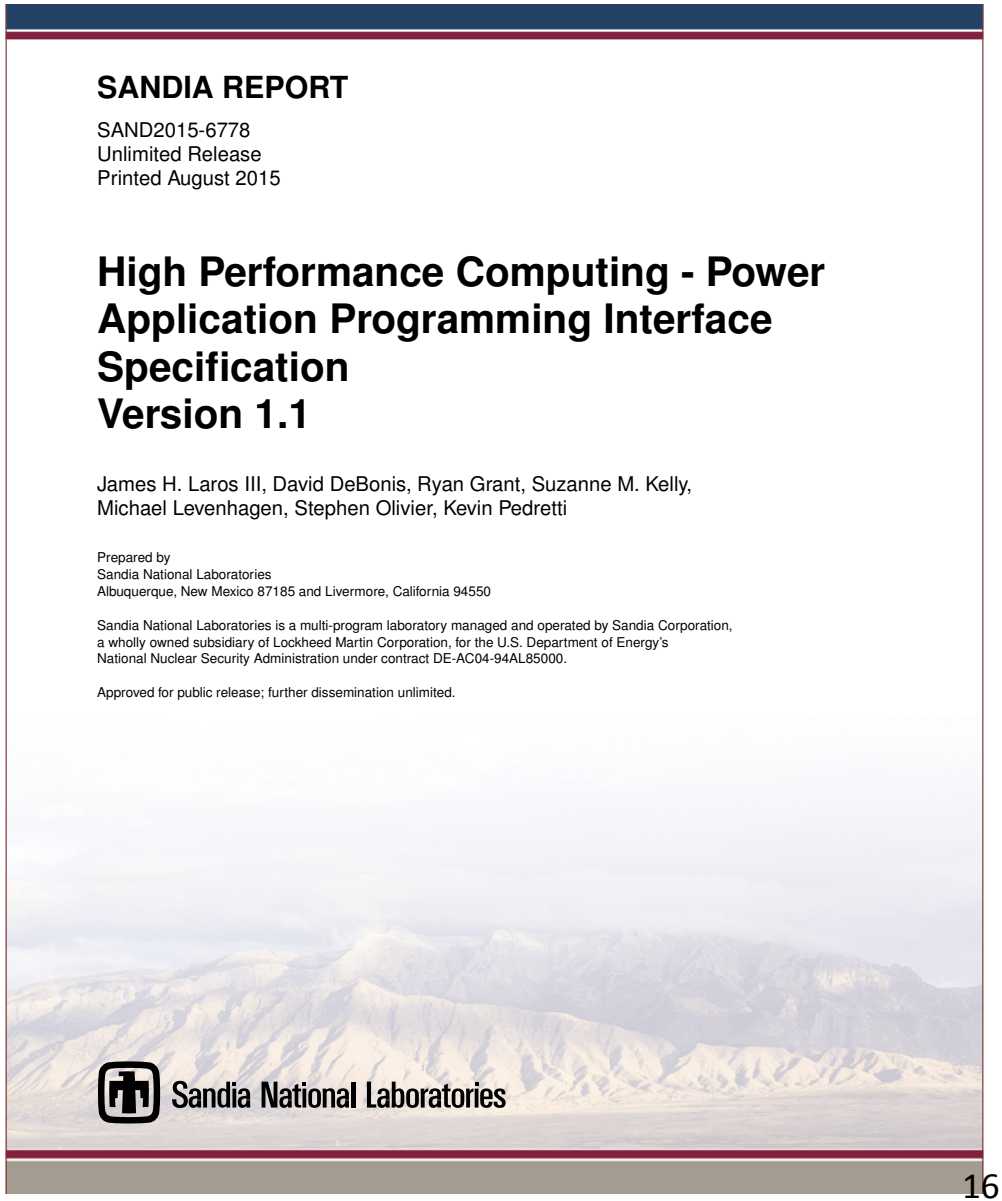
A Use Case Approach

HPC Power API Specification

Timeline Targets (FY2014)

- ✓ System Description development – end of December 2013
- ✓ Attribute Interface design – end of February 2013
- ✓ Early Draft ready for External Review – July 24th 2014
  - Small selected group of experts
  - Focused on Vendors (will be the implementers)
  - Face to Face attendance and written feedback required
  - Quite a lot to ask for a pizza lunch
    - **Cray, Adaptive, AMD, Intel, IBM, Penguin Computing, NREL**
    - A lot to ask for coffee and pizza
  - Additional written feedback
    - LANL, EEHPC, ORNL and UNM
  - ALL feedback addressed prior to public release
- ✓ First public release – August 26th 2014
- ✓ First public review – September 4th 2014 Denver Colorado
  - International audience
- Feedback: V 1.0 is good, more is needed

- Version 1.0 released August 2014
- Version 1.1 released August 2015
- Version 1.1a to be released this week
  - Minor corrections
  - Targeting Quarterly Releases
- Community needs a portable API for **measuring** and **controlling** power and energy
- Sandia developed Power API specification to fill this gap
- Provides portable power **measurement** and **control** interfaces, covers full spectrum of facility to component
- First production implementation will be Trinity (ATS-1)
  - Finalizing areas of focus with Cray
- Continued (increasing) community involvement and influence
- Reference Implementation on Sandia's Advanced Architecture Test Beds
- **To date 73 Countries (52% non-US)**
- http://powerapi.sandia.gov



SANDIA REPORT

SAND2015-6778
Unlimited Release
Printed August 2015

High Performance Computing - Power Application Programming Interface Specification
Version 1.1

James H. Laros III, David DeBonis, Ryan Grant, Suzanne M. Kelly, Michael Levenhagen, Stephen Olivier, Kevin Pedretti

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

System Description

PWR_ObjType

```
typedef enum {
        PWR_OBJ_PLATFORM = 0,
        PWR_OBJ_CABINET,
        PWR_OBJ_CHASSIS,
        PWR_OBJ_BOARD,
        PWR_OBJ_NODE,
        PWR_OBJ_SOCKET,
        PWR_OBJ_CORE,
        PWR_OBJ_POWER_PLANE,
        PWR_OBJ_MEM,
        PWR_OBJ_NIC,
        PWR_NUM_OBJ_TYPES,
        /* */
        PWR_OBJ_INVALID = -1,
        PWR_OBJ_NOT_SPECIFIED = -2
} PWR_ObjType;
```

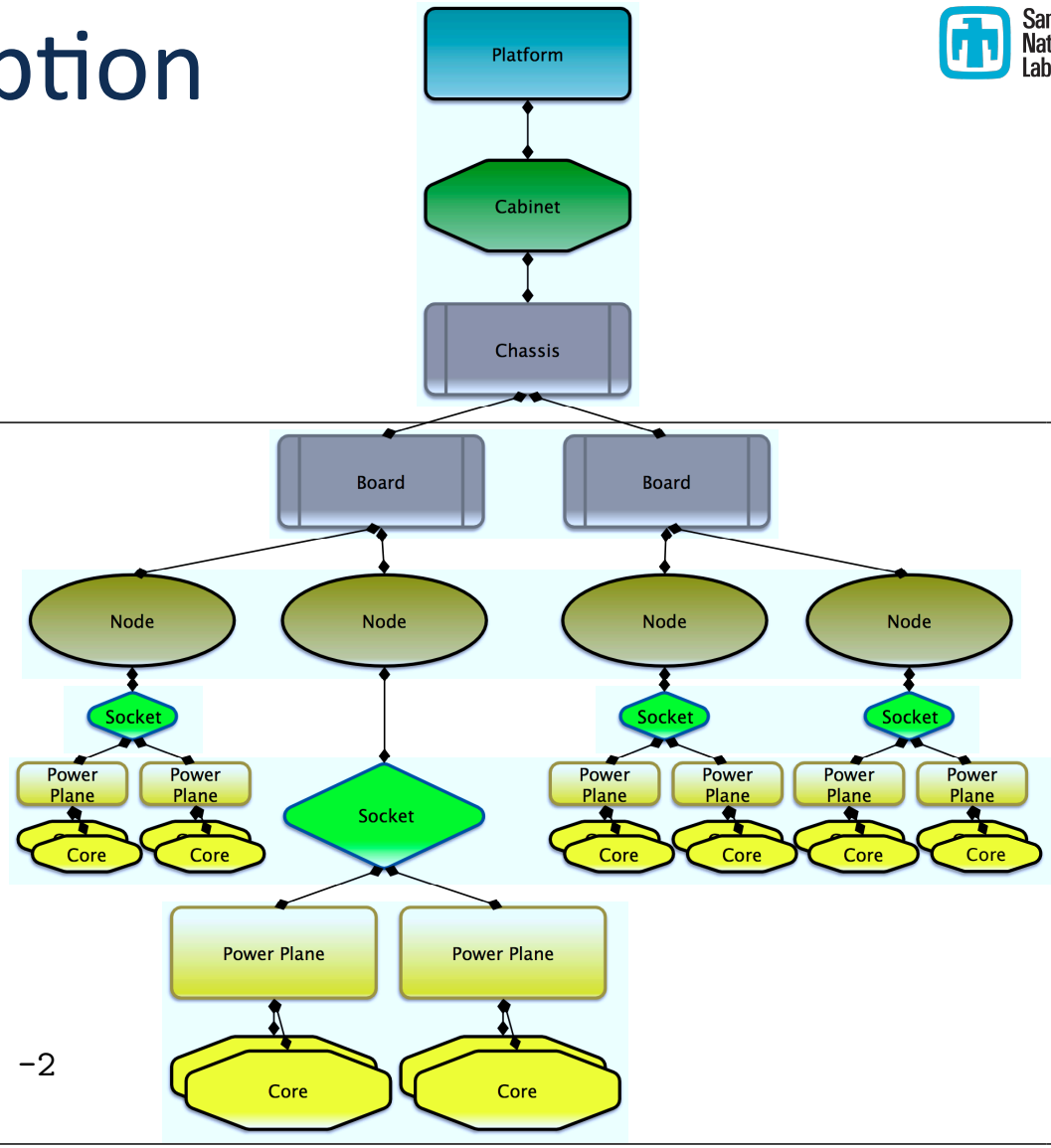
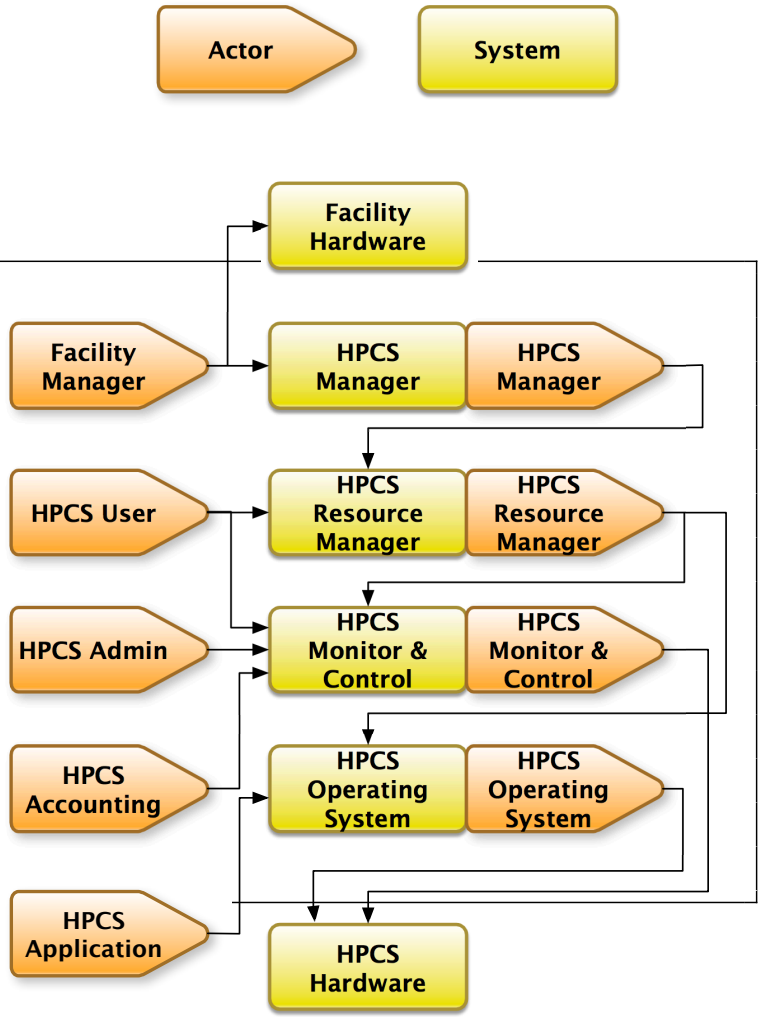


Figure 2.1: Hierarchical Depiction of System Objects

Roles

PWR_Role

```
typedef enum {
        PWR_ROLE_APP = 0, /* Application */
        PWR_ROLE_MC, /* Monitor and Control */
        PWR_ROLE_OS, /* Operating System */
        PWR_ROLE_USER, /* User */
        PWR_ROLE_RM, /* Resource Manager */
        PWR_ROLE_ADMIN, /* Administrator */
        PWR_ROLE_MGR, /* HPCS Manager */
        PWR_ROLE_ACC, /* Accounting */
        PWR_NUM_ROLES,
        /* */
        PWR_ROLE_INVALID = -1,
        PWR_ROLE_NOT_SPECIFIED = -2
} PWR_Role;
```



Foundation: Measurement and Control

PWR_AttrName

```
typedef enum {
        PWR_ATTR_PSTATE = 0, /* uint64_t */
        PWR_ATTR_CSTATE, /* uint64_t */
        PWR_ATTR_CSTATE_LIMIT, /* uint64_t */
        PWR_ATTR_SSTATE, /* uint64_t */
        PWR_ATTR_CURRENT, /* double, amps */
        PWR_ATTR_VOLTAGE, /* double, volts */
        PWR_ATTR_POWER, /* double, watts */
        PWR_ATTR_POWER_LIMIT_MIN, /* double, watts */
        PWR_ATTR_POWER_LIMIT_MAX, /* double, watts */
        PWR_ATTR_FREQ, /* double, Hz */
        PWR_ATTR_FREQ_LIMIT_MIN, /* double, Hz */
        PWR_ATTR_FREQ_LIMIT_MAX, /* double, Hz */
        PWR_ATTR_ENERGY, /* double, joules */
        PWR_ATTR_TEMP, /* double, degrees Celsius */
        PWR_ATTR_OS_ID, /* uint64_t */
        PWR_ATTR_THROTTLED_TIME, /* uint64_t */
        PWR_ATTR_THROTTLED_COUNT, /* uint64_t */
        PWR_NUM_ATTR_NAMES,
        /* */
        PWR_ATTR_INVALID = -1,
        PWR_ATTR_NOT_SPECIFIED = -2
} PWR_AttrName;
```

Attribute Interface

MEASURE

```
int PWR_ObjAttrGetValue( PWR_Obj object,
                         PWR_AttrName attr,
                         void* value,
                         PWR_Time* ts );
```

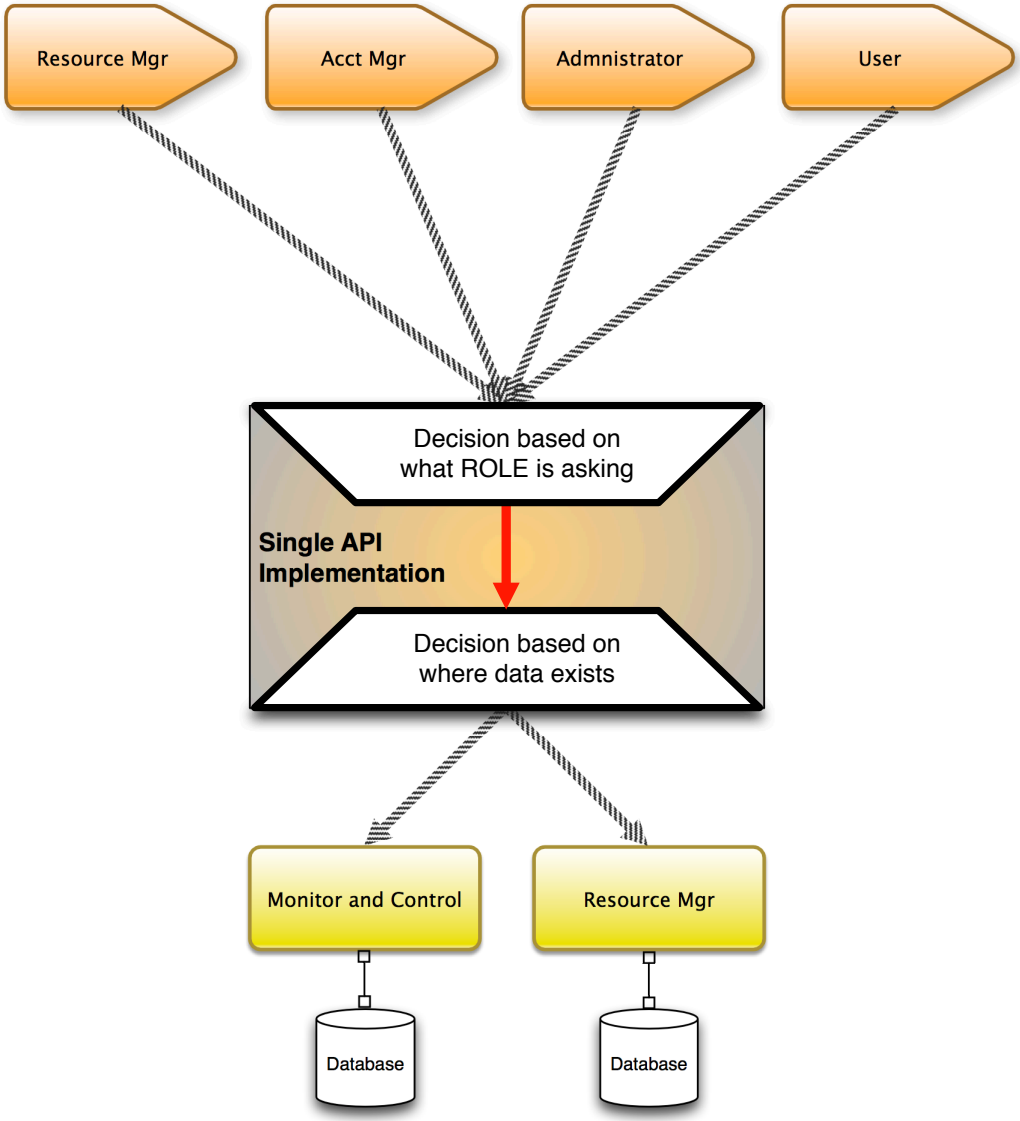
CONTROL

```
int PWR_ObjAttrSetValue( PWR_Obj object,
                         PWR_AttrName attr,
                         const void* value );
```

Other Core Features

- Discovery and Navigation
  - Of system description
- Groups
  - Adds lots of flexibility and convenience
- Statistics
  - Real-time, historic and extensible
- Metadata
- Languages
  - V 1.0 Specified in C
    - V 1.1 and 1.1a
  - V 1.x Python

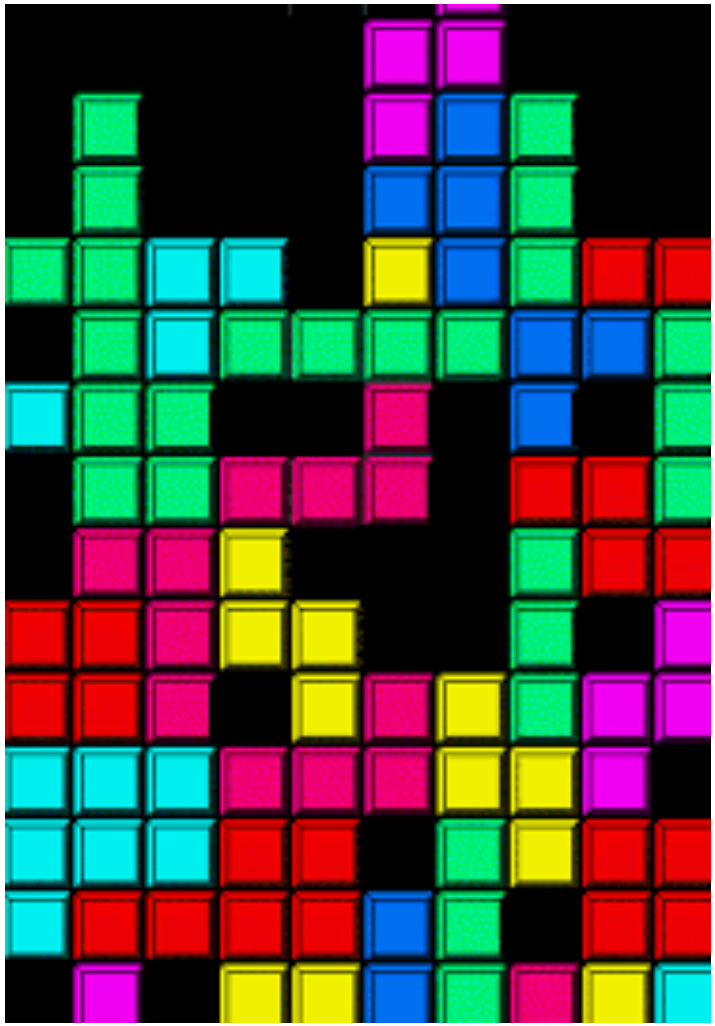
Higher Level Interfaces



Facility/Platform/ResMgr/App

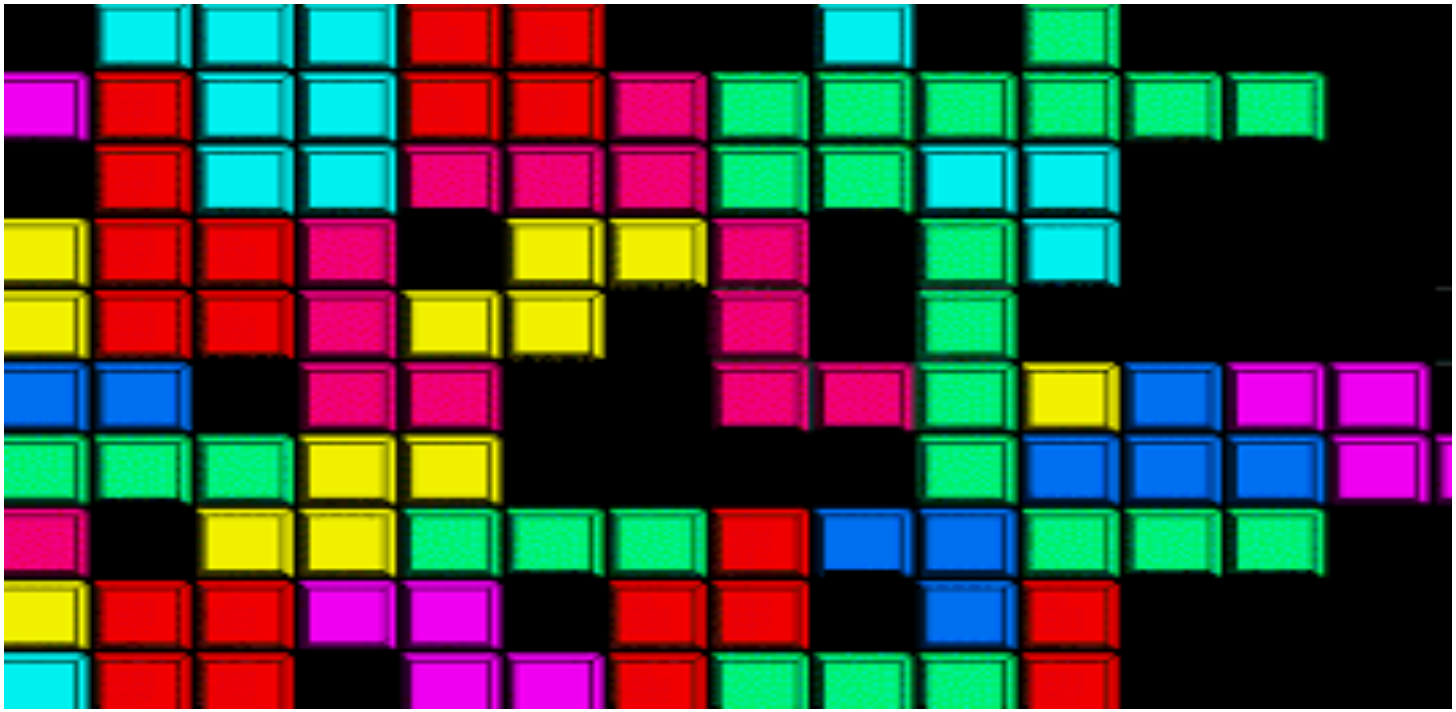
- Facility contracts with power provider
  - Based on anticipated need
    - Be great if we could guess better
  - Time of day implications
    - Day time premium
  - Violating contract may be largest cost motivation
- Manage platform within allocated budget
  - Sledgehammer approach, coarse hardware power caps
- Resource manager, per job allocations within platform boundaries
  - Best if this can be done intelligently
- Application, manage allocation to maximize efficiency (likely a mixture between throughput and power or energy)

Kind of like Tetris?



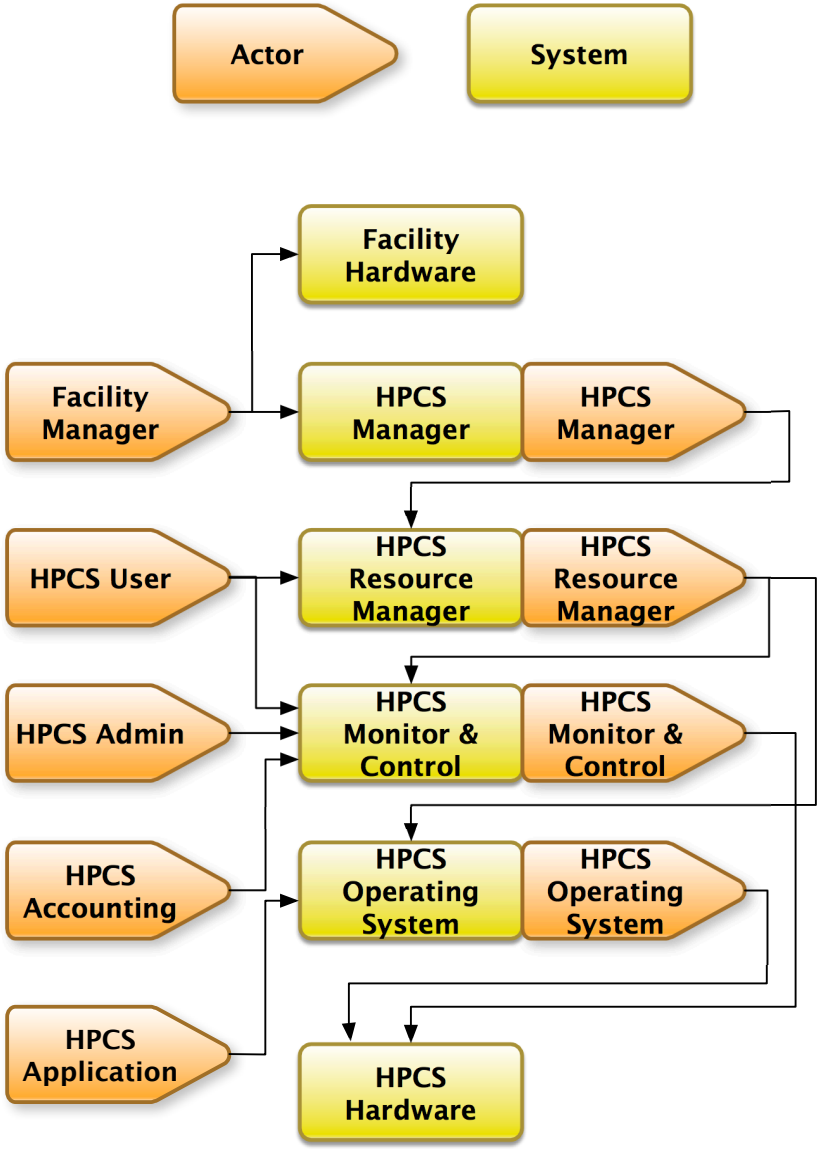
Kind of like Tetris?

POWER



TIME

Result


Actor
System
Facility Hardware
Facility Manager
HPCS Manager
HPCS Manager
HPCS User
HPCS Resource Manager
HPCS Resource Manager
HPCS Admin
HPCS Monitor & Control
HPCS Monitor & Control
HPCS Accounting
HPCS Operating System
HPCS Operating System
HPCS Application
HPCS Hardware

Questions?

- http://powerapi.sandia.gov
- Register on the reflector
- Get the current version of the spec
- Get the prototype and reference implementation source
- Other information as it develops
- Please get involved and help us improve the specification
- TEAM:

Suzanne Kelly, Kevin Pedretti, Michael Levenhagen, Ryan Grant, Stephen Olivier, David DeBonis