

Operation Pineapple: A Runtime Interposition Project



Joshua Hursey
hurseyjj@ornl.gov

Tuesday, June 12, 2012



U.S. DEPARTMENT OF
ENERGY



OAK RIDGE NATIONAL LABORATORY

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

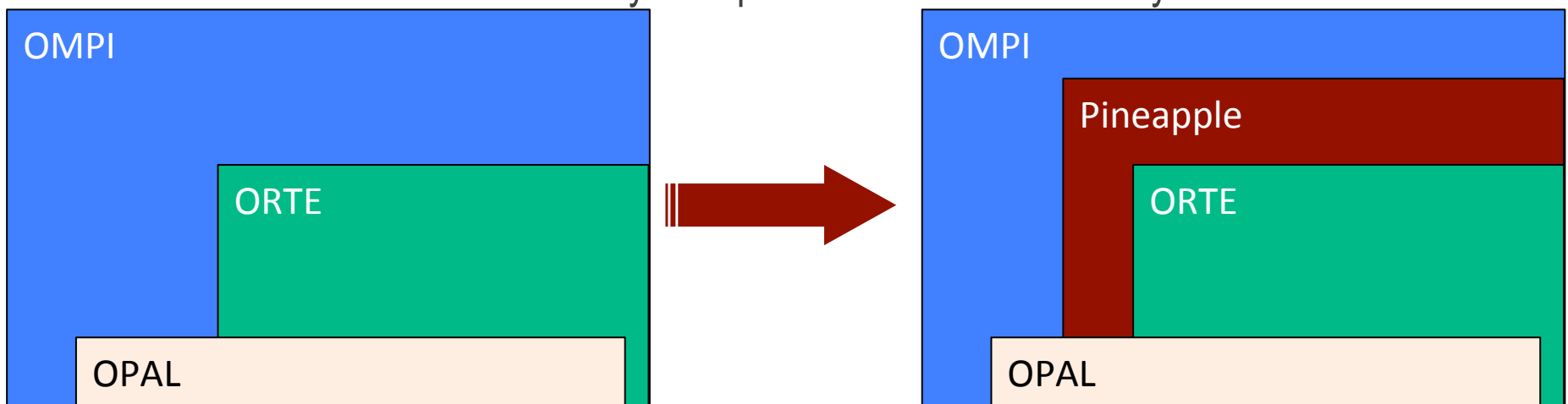
Why Pineapple?

A thorny issue with goodness inside.



Concept

- **Objective:** Encapsulate runtime requirements of OMPI behind a single API with well defined semantics.
- **Pineapple as a separate project in OMPI tree**
 - Neither 'owned' by OMPI nor ORTE
 - Exposes only the interfaces needed for OMPI (~ORTE Interface)
 - The **stems MCA framework** encapsulates the 'glue' code for RTEs
 - Protection from accidental symbol pollution into the OMPI layer

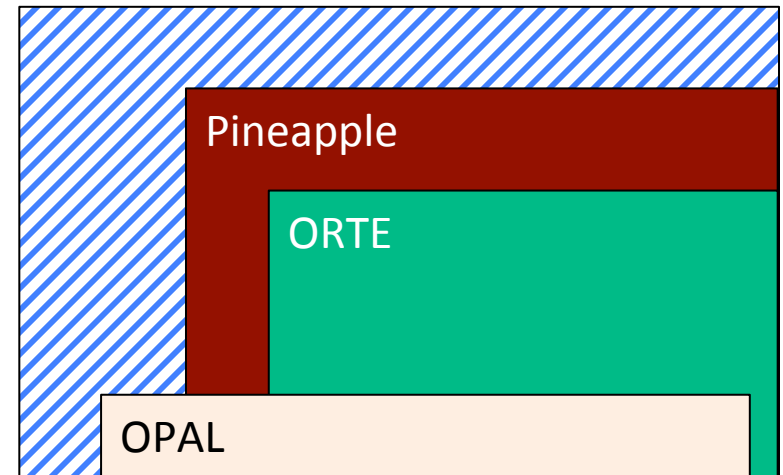


Concept: Neither 'owned' by OMPI nor ORTE

- Something other than (or in addition to) OMPI above ORTE

ORTE Owns

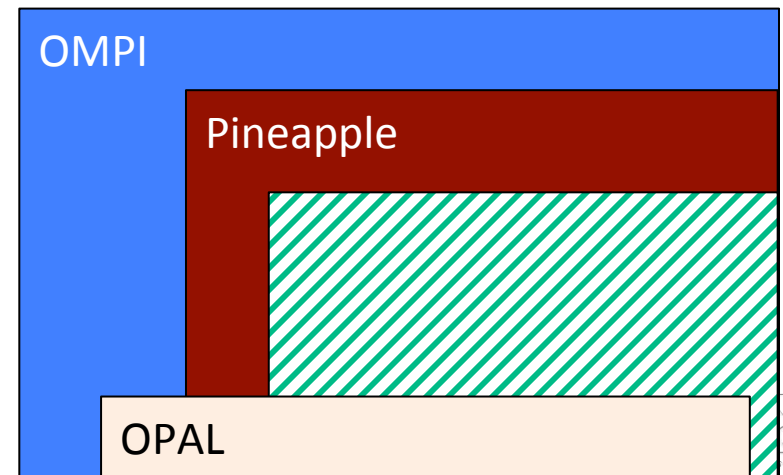
- UPC + OMPI (see MVAPICH)
- MapReduce
- Subset OMPI
- Research versions of OMPI
- Something other than MPI



- Something other than ORTE under OMPI

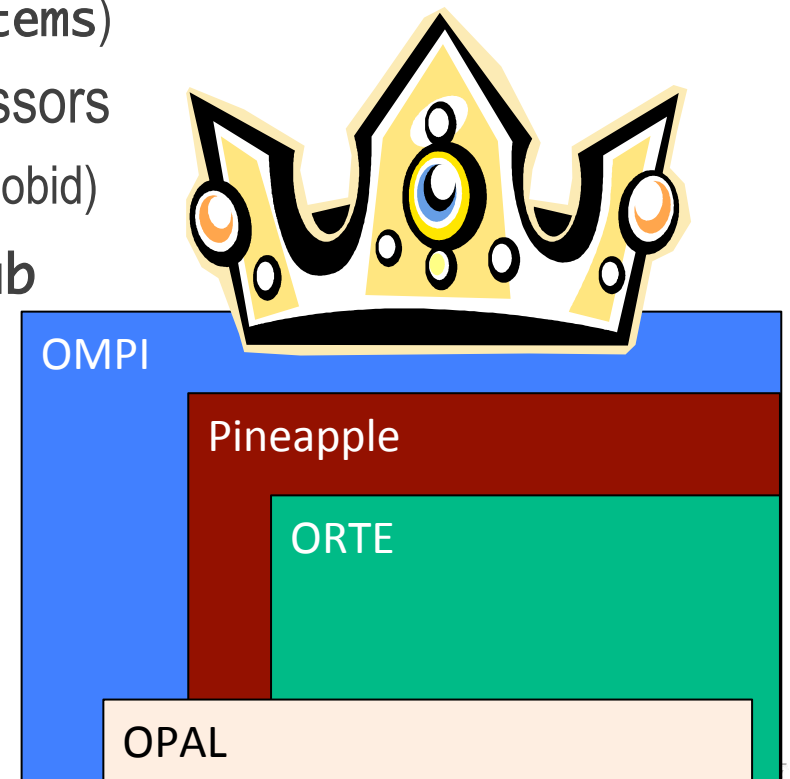
OMPI Owns

- 'minimal' component (Sandia)
- Research versions of ORTE
- Other experimental research RTEs
- System specific RTEs



Interface

- Interfaces match subset of ORTE currently used by OMPI
 - `#include "pineapple/include/rte_pineapple.h"`
 - Crisp delineation of what OMPI needs from an RTE
 - Clear definition of semantics in the header
 - Framework to abstract RTE specific code (**stems**)
 - Data structures are hidden behind accessors
 - Process name – static inline functions (vpid, jobid)
 - OMPI layer frameworks **dpm** and **pubsub**
 - Should be moved to the pineapple layer
- **OMPI/ORTE/OPAL Stack is King!**
 - Pineapple interface should always err on the side of supporting the default OMPI/ORTE/OPAL stack

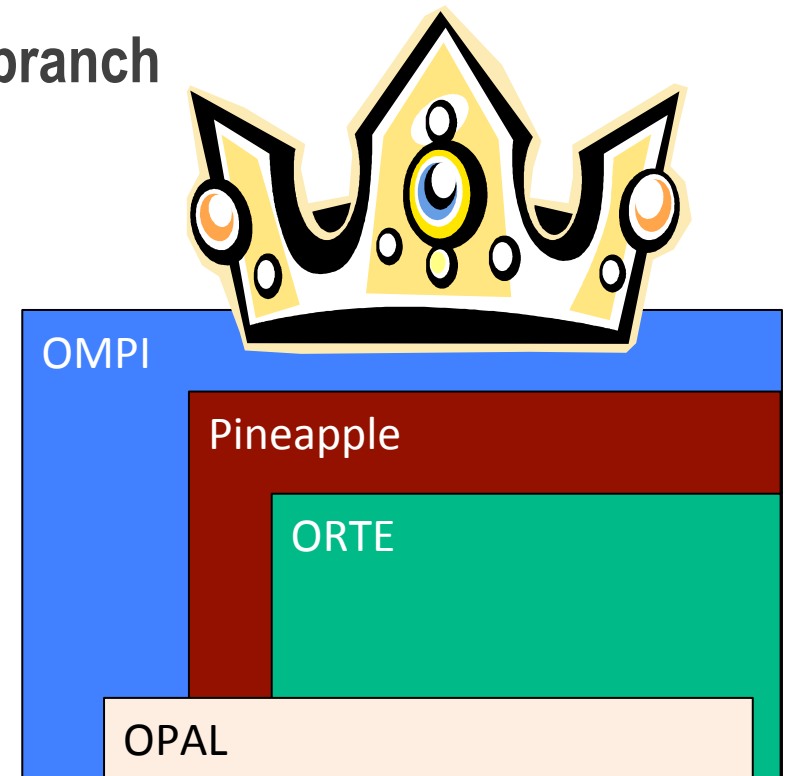


Todo Items

- **Todo Items:**
 - **OMPI-level tools that depend on ORTE**
 - **ompi_info:** If !ORTE then do not build
 - Do something better down the road
 - **ortetools:** If !ORTE then do not build (link)
 - **ompi-server:** If !ORTE then do not build
 - **Finish converting remaining components of OMPI**
 - Move dpm and pubsub to pineapple
 - **Testing, testing, testing...**
 - **Name: "O+{3 letters}" (e.g., OMPI, ORTE, OPAL)**
 - http://www.wordbyletter.com/words_starting_with.php?q=o&letters=4
 - <http://www.morewords.com/wordsbylength/4o/>
 - OATH, OBEY, OGLE, OKAY, OKRA, OMEN, OMIT, OOZY, OPAH, ORCA, OUCH, ...

Timeline

- Targeting 1.7 release series
- **Timeline: 2 weeks** (Last week of June)
 - M 6/25: Finish development in the branch
 - T 6/26: Discuss on teleconf
 - R 6/28: Commit to trunk
 - M 7/16: Push to 1.7trunk



<https://bitbucket.org/jjhursey/ompi-pineapple>