



Программные методы контроля и диагностики вычислительных систем в системах управления ресурсами



Сибирский государственный университет телекоммуникаций и информатики,
кафедра Вычислительных систем

Институт физики полупроводников им. А.В. Ржанова,
лаборатория Вычислительных систем

Крамаренко Константин Евгеньевич
kostya.kram@gmail.com

Тринадцатая международная азиатская школа-семинар "Проблемы оптимизации сложных систем" в рамках
международной мультikonференции IEEE SIBIRCON 2017
18-22 сентября 2017, Новосибирск

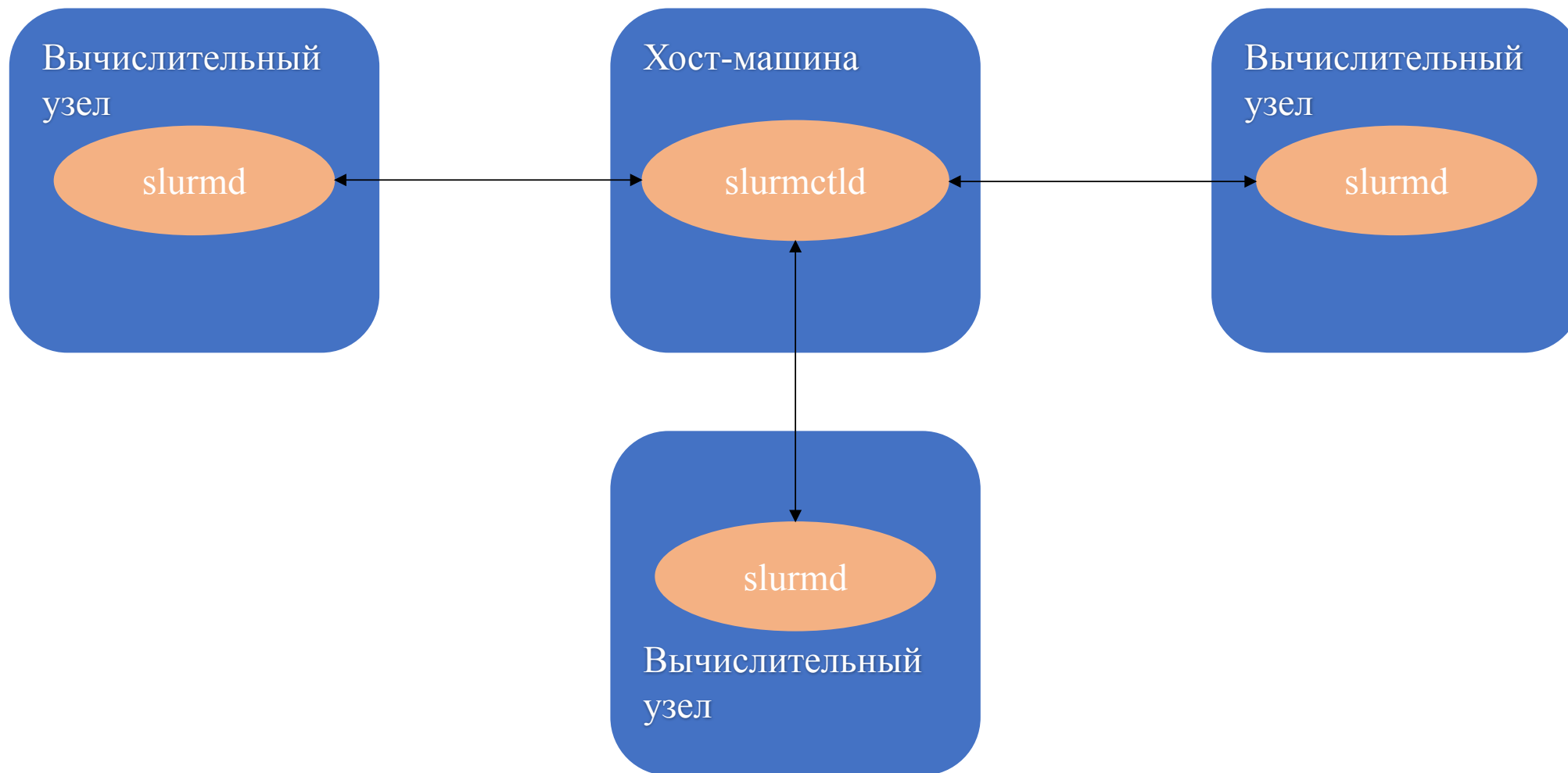
Системы управления ресурсами ВС



Torque

[*] Список Top500, редакция июня 2017г. - <https://www.top500.org/lists/2017/06/>

Архитектура SLURM



Алгоритмы контроля и диагностики состояния узлов в системе SLURM

```
function SLURMCTLDBACKGROUND()  
  while TRUE do  
    if (TIME() – LastPingTime >= PingInterval) then  
      LastPingTime = TIME()  
      PINGNODES()  
    end if  
    if (TIME()– LastHealthCheckTime >= HealthCheckInterval) then  
      LastHealthCheckTime = TIME()  
      RUNHEALTHCHECK()  
    end if  
  end while  
end function
```

$$T_{SlurmctldBackground} = O\left(T_{PingNodes} + T_{RunHealthCheck}\right).$$

Алгоритмы контроля и диагностики состояния узлов в системе SLURM

```
function PINGNODES()  
    PingNodeList = LISTCREATE()  
    RegNodeList = LISTCREATE()  
    for each Node in NodeList do  
        if (ISNODENOTREGISTERED(Node)) then  
            LISTPUSH(RegNodeList, Node)  
            continue  
        end if  
        if (ISNODENORESPOND(Node) OR ISNODEDOWN(Node)) then  
            continue  
        end if  
        LISTPUSH(PingNodeList, Node)  
    end for  
    SENDREQUEST(PingNodeList)  
    SENDREQUEST(RegNodeList)  
end function
```

$$T_{PingNodes} = O\left(n + T_{SendRequest}\right),$$

n – количество узлов в ВС,

$T_{SendRequest}$ – время рассылки служебной информации с применением древовидного алгоритма, имеет порядок $O(\log n)$.

Алгоритмы контроля и диагностики состояния узлов в системе SLURM

```
function RUNHEALTHCHECK()  
    HealthCheckNodeList = LISTCREATE()  
    for each Node in NodeList do  
        if (ISNODENORESPOND(Node) OR ISNODEDOWN(Node)) then  
            continue  
        end if  
        LISTPUSH(HealthCheckNodeList, Node)  
    end for  
    SENDREQUEST(HealthCheckNodeList)  
end function
```

$$T_{RunHealthCheck} = O\left(n + T_{SendRequest}\right),$$

n – количество узлов в ВС,

$T_{SendRequest}$ – время рассылки служебной информации с применением древовидного алгоритма, имеет порядок $O(\log n)$.

Заключение

- Распараллеливание основных фаз алгоритмов контроля и диагностики
- Применение децентрализованных алгоритмов
- Реализовать на уровне систем параллельного программирования (MPI+ULFM)

Спасибо за внимание!

Крамаренко Константин Евгеньевич
kostya.kram@gmail.com

Институт физики полупроводников им. А.В. Ржанова,
лаборатория Вычислительных систем

Сибирский государственный университет телекоммуникаций и информатики,
кафедра Вычислительных систем