Virtualisation
M2 – P7
2015

François Armand

Université Paris Diderot

A enda !1/2"

- Use #ases
- Multi-#ore #$allen es
- Virtuali%ation &a'ono ( )
- Virtuali%ation* $o+ ,
  - &rans-arent virtuali%ation*
    - .ard+are assisted virtuali%ation
    - D)na (i# /inar) &ranslation
  - Para-virtuali%ation
- Virtuali%ation and Devi#es

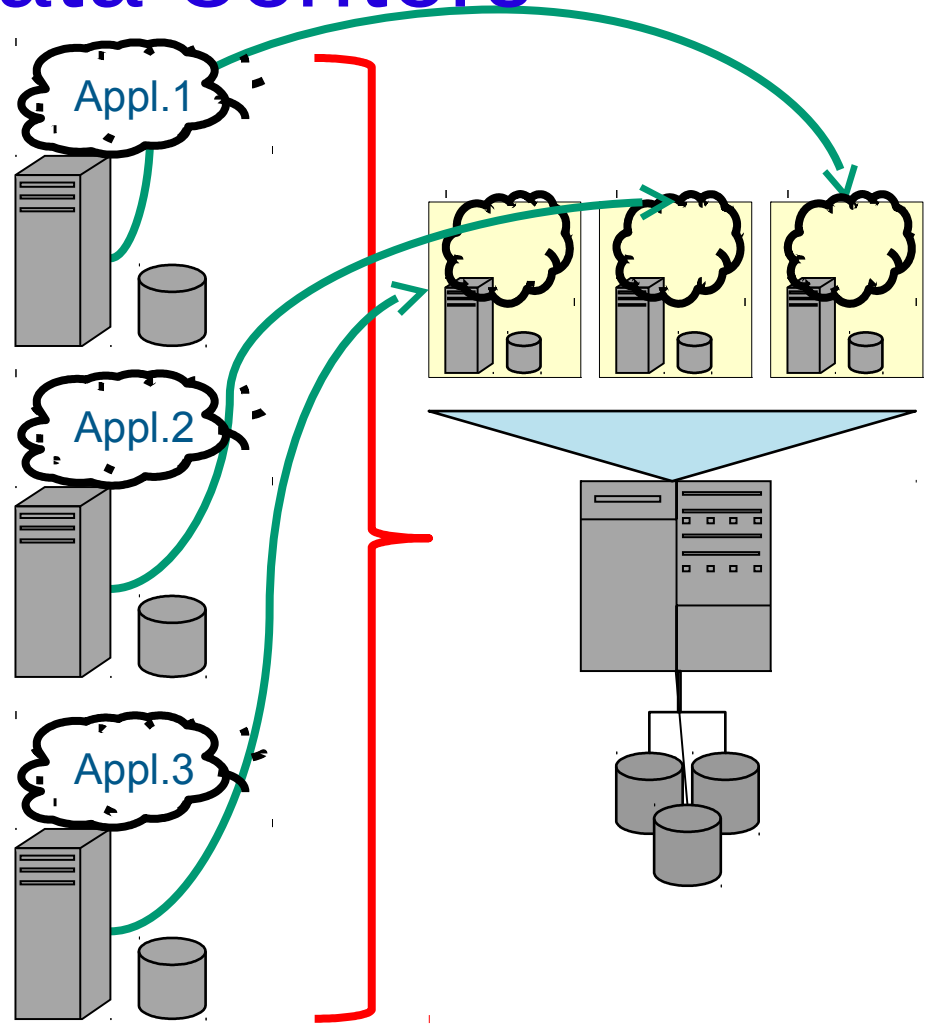
A enda !2/2"

- . ard+are 0volution
- . osted versus 1ative Virtuli%ation
  - 0'a (-les * 2VM3 4en3 0S4
- Advan#ed Virtuali%ation #a-a5ilities

- I (-a#t on 6S and a--li#ations -a#7a in
- Standardi%ation

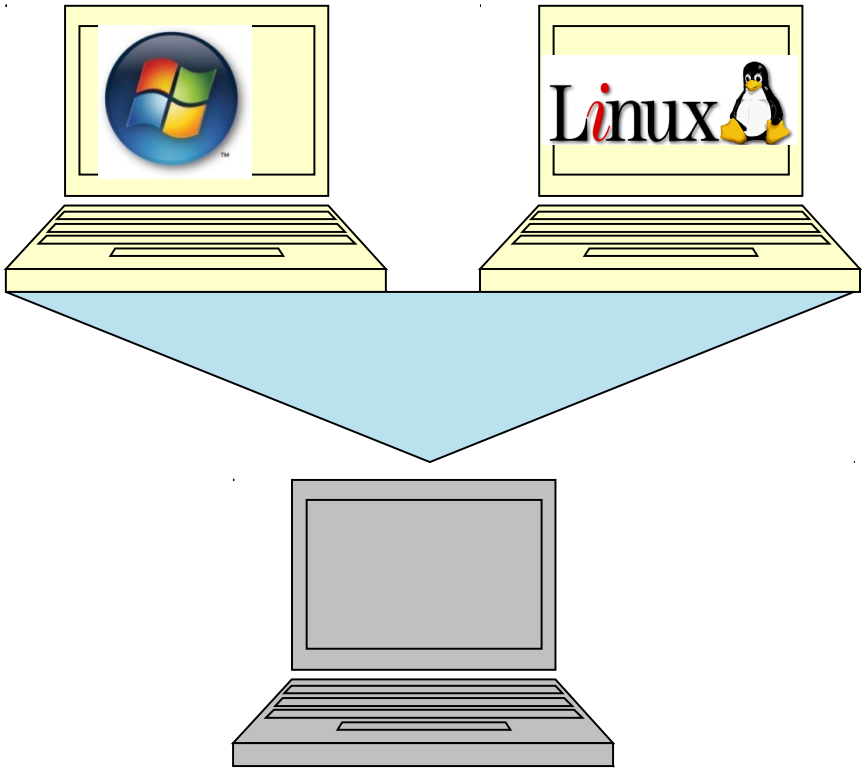
Use 8ase* Data 8enters

- 8onsolidate +or7loads runnin on inde-endent -$)si#al (a#$ines on a sin le server3 +$ile (aintainin inde-enden#e
- 9Sto- Ma#$ine s-ra+l: ☺
- S-lit !virtuali%e" -$)si#al servers into 9s (aller: Virtual Ma#$ines
- Allo#ate VM d)na (i#all)

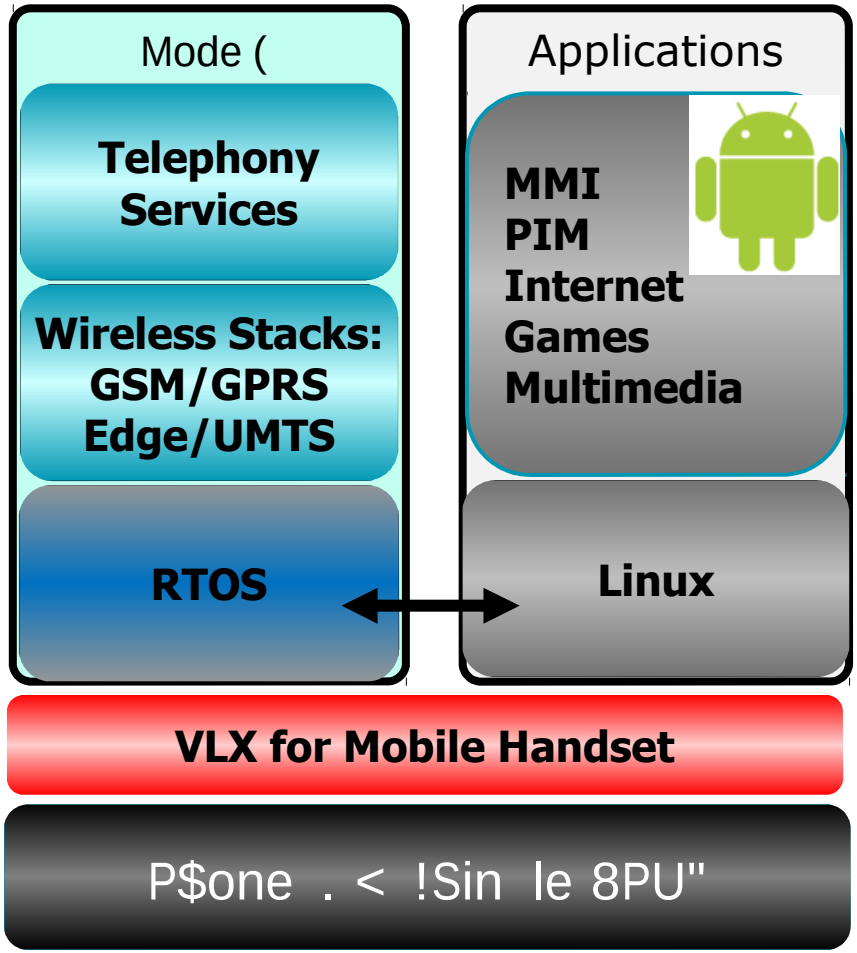


Use 8ase* <or7station

- Si (ilar to data #enters
- Run (ore t$an a sin le environ (ent at on#e on a sin le (a#$ine



Use 8ase* Mo5ile . andsets

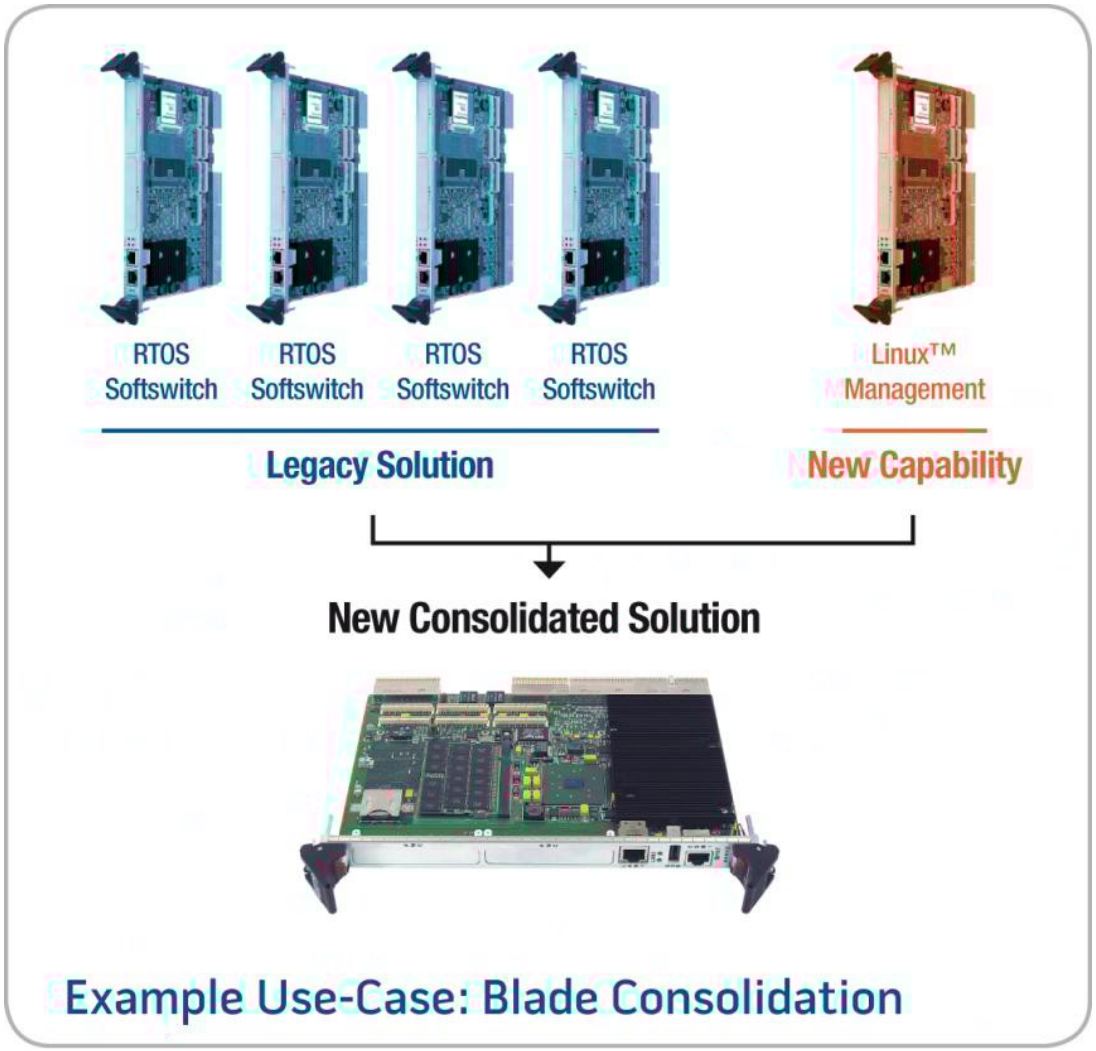


- Run =inu' a--li#ations on 5ase5and -ro#essor
- Re-use e'istin (ode (so›t+are sta#7 +it$ its R&6S !no #$an es"
- Su--ort o› =inu' at a (ini (al develo- (ent #ost
- 6-eratin S)ste (inde-enden#e ›or ›uture evolutions
- See also 1&& Do#o (o 6S&I

Use 8ase* Mo5ile . andsets

- Virtual=o i' V=4 *! - revious slide"*
  - 14P -lat›or (3 2 A#er (o5ile -$ones
- VM+are Mo5ile Virtual Plat›or (!MVP"
  - &ran o Virtual Pro#essor a#?uisition
- 6 -en 2ernel =a5
  - =; Mi#rovisor
  - Used 5) @ual#o ((and ot$ers ! . &83 MotorolaA"

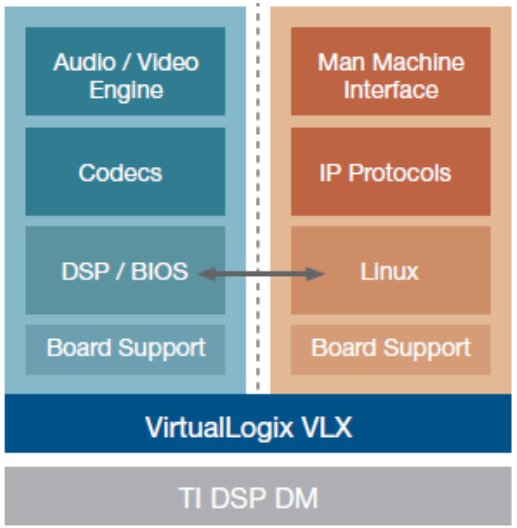
Virtuali%ation in .i $-&$rou $-ut 1et+or7 0?ui- (ent



Use #ase*

Virtuali%ation in Multi (edia Devi#es

- Redu#tion o> / 6M >or $i $-volu (e lo+-end -rodu#ts
  - 1o need >or a Beneral Pur-ose Pro#essor
    - C 20 to 25 D / 6M redu#tion
  - Run =inu' to et$er +it$ 6S su--ortin 8ode#s on a sin le &I DSP
  - =evera e =inu' environ (ent
  - Reuse e'istin DSP so>t+are



6t$er Use 8ases

- Instru (entation3 Auto (ation
  - Run a R&6S and a BP6S >or Bra-$i#al Inter>a#e
- Points o> Sales
  - Run t$e UI and t$e se#ure transa#tion environ (ent on t$e sa (e -ro#essor
- Mil / Aero
  - Run se#urel) isolated / #erti>ied environ (ents si (ultaneousl)
- MoreA

A enda !1/2"

- Use #ases
- Multi-#ore #$allen es
- Virtuali%ation &a'ono ( )
  - S)ste (level versus Pro#ess level virtuali%ation
  - 1ative versus . osted Virtuali%ation
  - &rans-arent virtuali%ation versus -ara-virtuali%ation
- 0 (5edded E Real-&i (e Virtuali%ation Re?uire (ents
- . ard+are 0volution

Multi-8ore 8$allen es

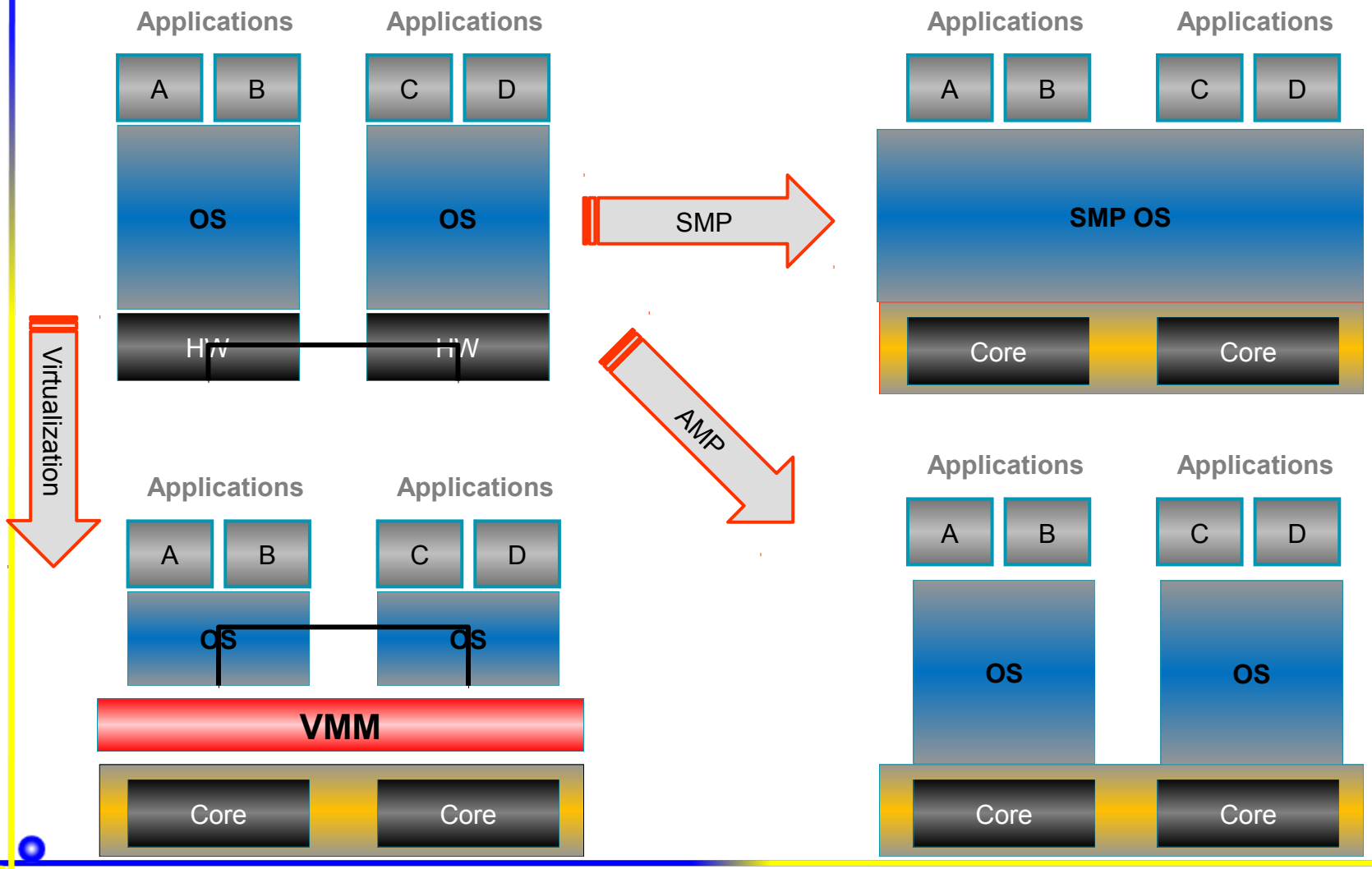
- In#reased -o+er no lon er -rovided 5) s-eed in#rease3 5ut t$rou $ $i $er densit) and (ulti-#ore #$i-s
- Man) le a#) R&6S $ave 5een desi ned 5ased on uni--ro#essor assu (-tion
  - S#alin on (ulti-#ore re?uires ti (e
  - .as 5een !is still" a -ain >or BP6S !s#ala5ilit) issues >or 6S and A--li#ations"
- =e a#) R& a--li#ations desi ned >or uni--ro#essor as +ellA 1eed to (ove to true (ulti-t$read +it$ true -arallelis (AF
  - 0ven (ore di>>i#ult t$an >or 6S%s

Multi-8ore 8$allen es

- Virtuali%ation ena5les to run (ulti-le instan#es o> a H*le a#) real-ti (e*l so>t+are load Jdesi ned >or uni--ro#essor- si (ultaneousl) on a (ulti-#ore #$i-
- Possi5ilit) to run in uni--ro#essor Virtual Ma#$ines3 assi ned to a -$)si#al 8PU
- For BP6S* run n-+a) SMP instan#es on a lar er (a#$ine !e'* ; ;-+a) =inu' on a 1K #ores (a#$ine"
- 1o (odi>i#ations o> le a#) #ode3 1o ris7 to 5rea7 #ode
- S#ala5ilit) -rovided 5) Virtuali%ation la)er

Virtuali%ation 0na5les 8onsolidation

- &a7in advanta e o> (ore -o+er>ul $ard+are



A enda !1/2"

- Use #ases
- Multi-#ore #$allen es
- Virtuali%ation &a'ono ( )
  - S)ste (level3 Pro#ess level virtuali%ation3 6S virtuali%ation
  - 1ative versus . osted Virtuali%ation
  - &rans-arent virtuali%ation versus -ara-virtuali%ation
- 0 (5edded E Real-&i (e Virtuali%ation Re?uire (ents
- . ard+are 0volution

Virtuali%ation i4 +as ettin interestL

Virtuali%ation, 6$3 Virtuali%ationL

Virtual Networking?

Intel VT?

UML?

IBM/VM?

Virtual Server?

VMware?

Transitive QuickTransit?

Java, JVM?

Application Virtualization?

Pascal Pcode?

Platespin?

AMD SVM?

Virtual Solutions?

TransMeta Crusoe?

SIMICS?

Softricity?

QEMU?

KVM?

Dynamo?

VirtualPC?

Virtual Reality?

Virtual Storage?

FX!32?

Di>>erent #lasses o> virtuali%ation

- A re ation
  - Ma7e 1 resour#es a--ear as 1 !#lusters / rids3 dis7s"
- Partition / Re-li#ation *
  - Ma7e a resour#e a--ear as 1 !VMs3 dis7s"
  - Ma) 5e #o (5ined +it$ 9a re ation: !dis7s"
- &ranslation !0 (ulation"
  - Ma7e 4 a--ear as M *!so (eti (es 4 is identi#al to M"*
  - Ma) 5e #o (5ined +it$ 9-artition:

Mostl) interested in 9-artition:

. istor)

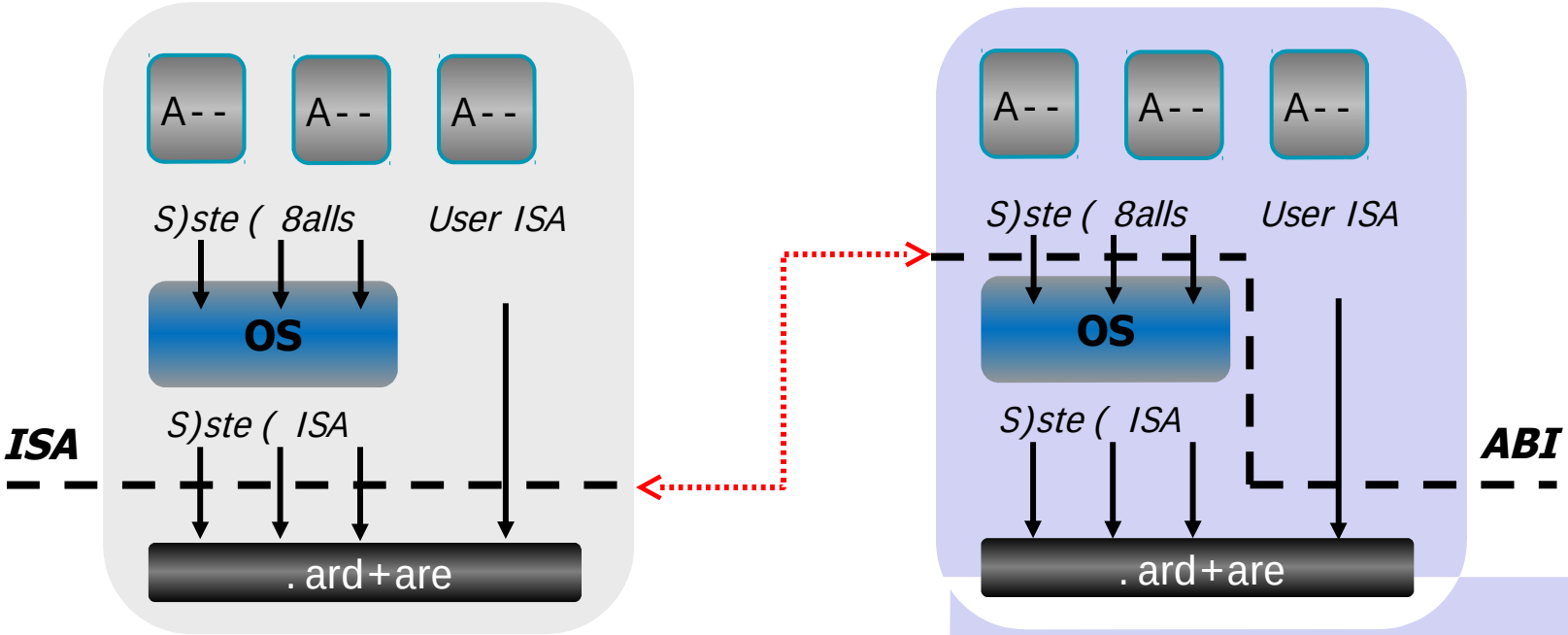
- A--eared durin t$e K0Gs3 I/M
- Po-ular to-i# durin t$e K0Gs and t$e 70Gs
- 9Surve) o> Virtual Ma#$ine Resear#$: !Bold5er 3 1N7;"
  - 9Virtual Ma#$ines $ave >inall) arrivedF Dis (issed >or a nu (5er o>)ears as (erel) a#ade (i# #uriosities3 t$e) are no+ seen as #ost-e>>e#tive te#$ni?ues >or or ani%in #o (-uter s)ste (s resour#es to -rovide e'traordinar) s)ste (>le'i5ilit) and su--ort >or #ertain uni?ue a--li#ations:
- 72 -a-ers -u5lis$ed >ro (1NKK to 1N70

. istor)

- De›inition 5) Bold5er !K0-70" *
- Si (ulation o› t$e instru#tion set 4 on usin t$e instru#tion set B
  - Slo+do+n 1000 B instru#tions -our 1 4 instru#tion
- <$en 4 PP B
  - Slo+do+n * 20 -our 1
- <$en 4 PP B3 use B as (u#$ as -ossi5le ›or 5etter e››i#ien#)
  - Virtual Ma#$ine s)ste (s !VMs"
  - Si (ulator P Virtual (a#$ine (onitors !VMM" (ostl)
- 1otion o› 9Full) Virtuali%a5le Pro#essors:

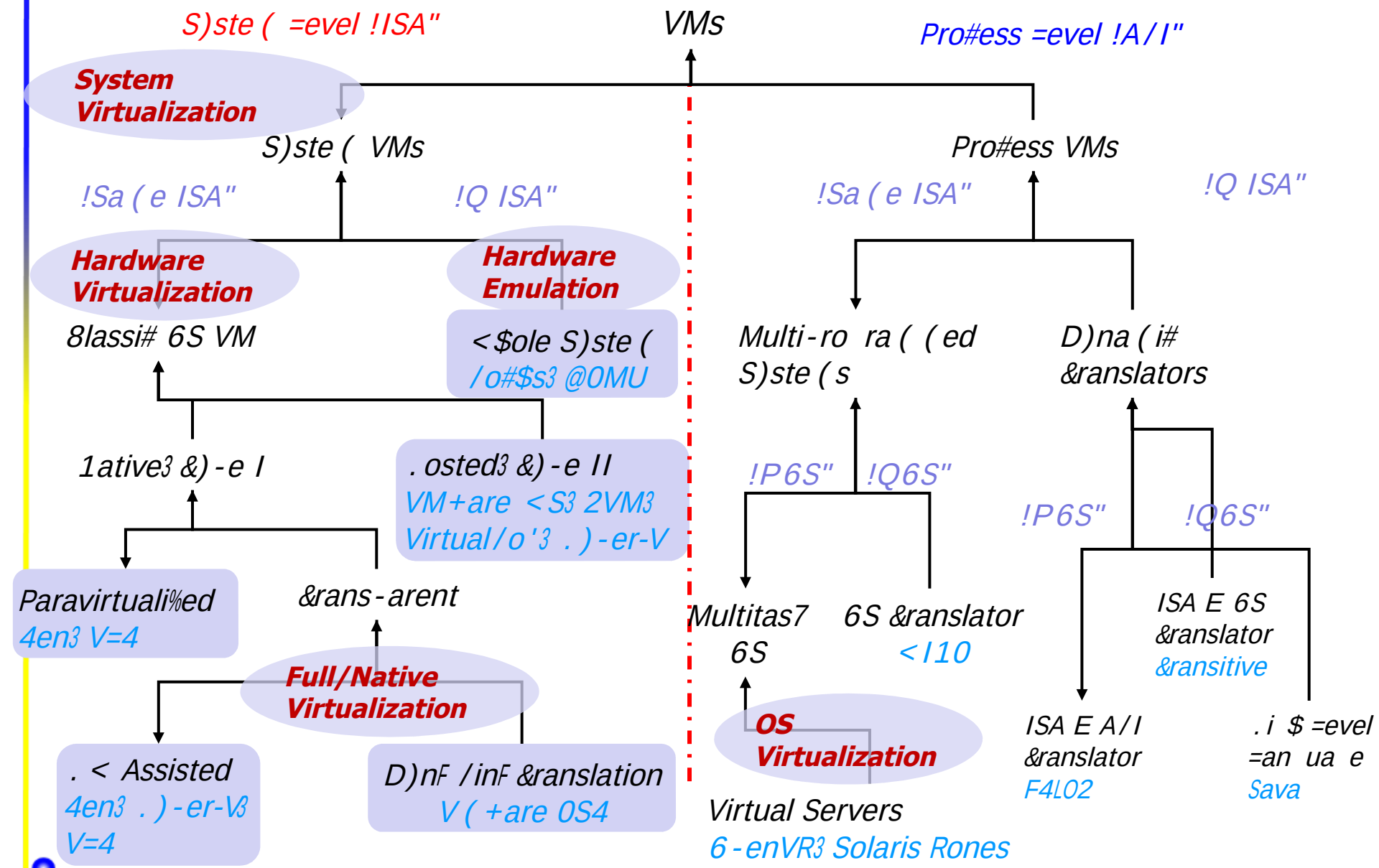
Virtual Ma#$ine Inter>a#es

- S)ste (s 5uilt o> * $ard+are3 6S3 a- -li#ations
- 2 (ain inter>a#es* ISA !$ard+are"3 A/I !6S"

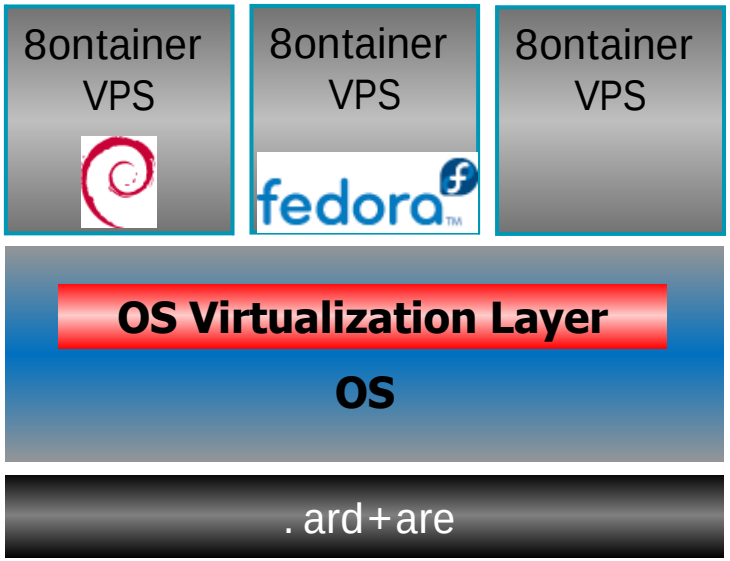


- S)ste (level VM
  -rovide an ISA inter>a#e
- Pro#ess level VM
  -rovide an A/I inter>a#e

&a'ono ( ) !derived >ro (OF S (it$ E 1air"



6S Virtuali%ation or Virtual Servers



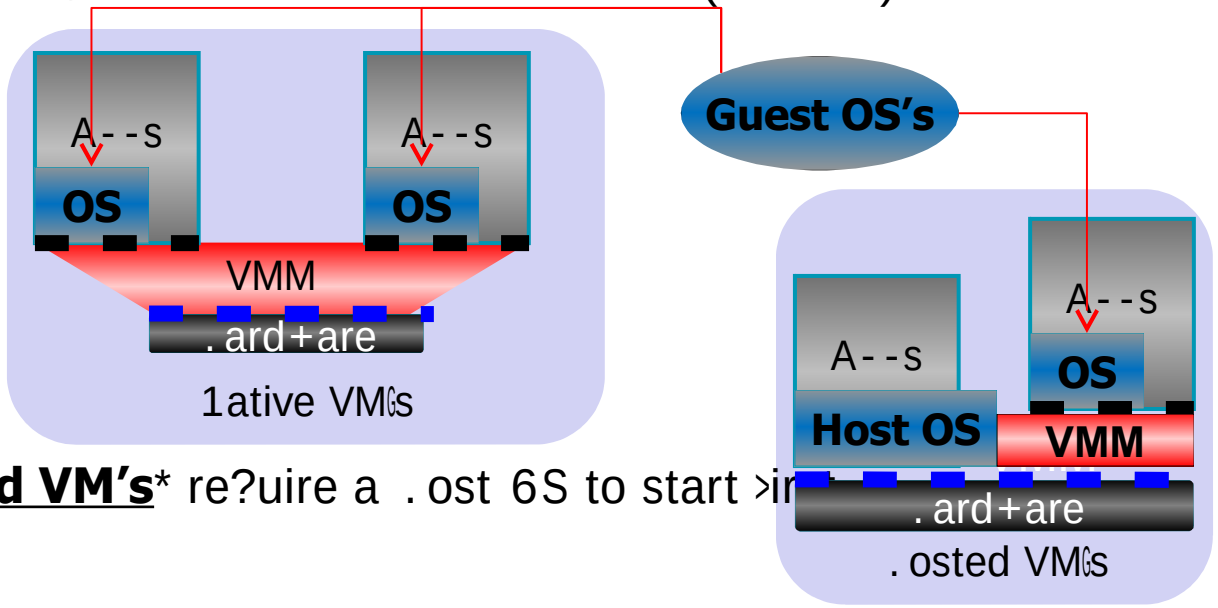
- 0'a (-le*
  - Virtuo%%o / 6-enVR
- Pro%s*
  - Su--orts distri5ution $etero eneit)
  - =i $t+ei $t*
    - lo+ (e (or) and -er>or (an#e over$ead
  - S#ales to (an) instan#es
- 8on%s*
  - Sin le 6S instan#e !#o ((on -oint o> >ailure"
  - Modi>ied 6S3 intrusive* $as to >ollo+ 6S evolution
  - Does not su--ort 6S $etero eneit) !. en#e no R&6S / BP6S #o (5ination"

6S Virtuali%ation * 0'a (-les

- AI4 <or7in Partitions
- Solaris Rones3 8ontainers
- Free/SD Tail
- =inu'-Vserver3 FreeVPS
- Virtuo%%o / 6-enVR
- i8ore Virtual A##ounts !<indo+s" ,

8lassi# 6S VMs

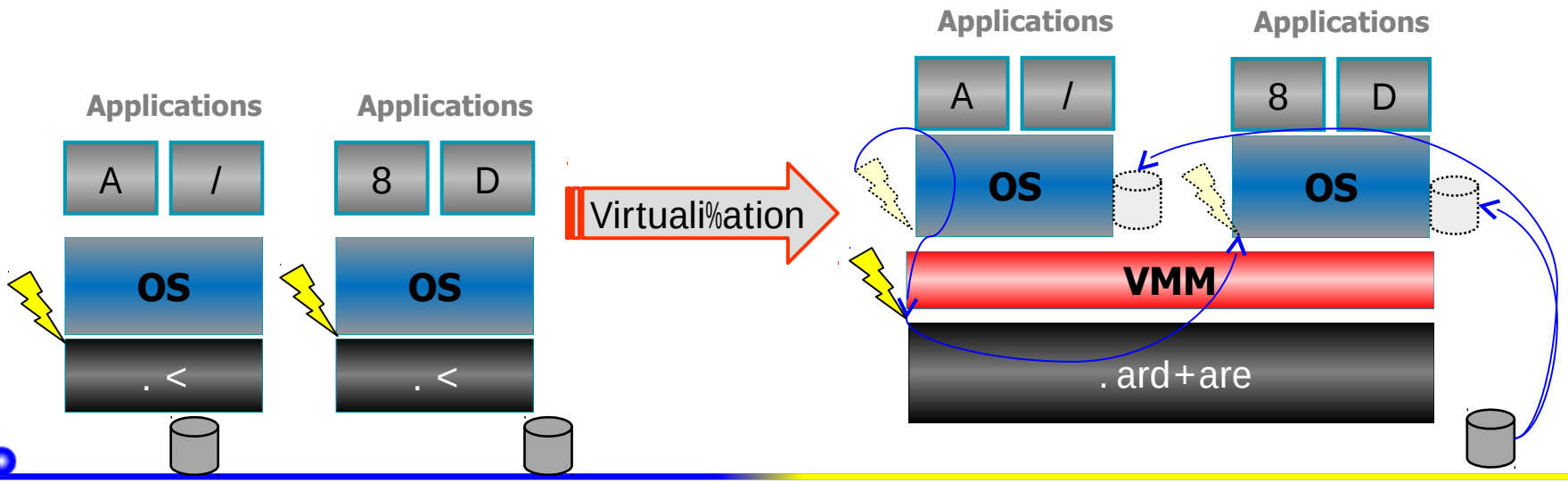
- 0na5le to run (ulti-le Jinde-endent- 6S%s si (ultaneousl) on t$e sa (e -ro#essor!**guest os**"3 ea#$ in its o+n **"Virtual Machine**:
- &+o (ain a--roa#$es*
  - **Native VM's*** Introdu#e a so>t+are la)er 5et+een t$e $ard+are and t$e 6S* Virtual Ma#$ine Monitor !VMM" or 9/are (etal: .)-ervisor



  - **Hosted VM's*** re?uire a .ost 6S to start >ir

8lassi# 6S VMs * Issues

- Run (ulti-le 6S%s
  - 6S desi ned assu (in t$e) are t$e onl) so>t+are #ontrollin t$e -$)si#al resour#es o> t$e (a#$ineF
- 8on#urrent a##ess to -$)si#al resour#es *
  - 8PU PU s#$eduler3 Me (or) PU -artition3 MMU , Devi#es ,
- 1eed to dete#t and resolve #on>li#ts su#$ as
  - (as7in interru-ts3 initiatin an I/63 or -ro ra((in t$e MMU
- Provide t$e e'-e#ted 5e$avior +it$in t$e Virtual Ma#$ineF



8lassi# 6S VMs * Issues

- Dete#t sensitive instru#tions and 9>a7e: t$e (F
- &rans-arent +a)s
  - Bood $ard+are su--ort*
    - 0'e#ute 6S at lo+er $ard+are -rivile e !iFeF * user (ode" to tra- su#$ instru#tions u-on e'e#ution
  - <ea7 $ard+are su--ort*
    - 8ir#u (vent non tra--in instru#tions
- 1on &rans-arent +a)
  - Modi>) 6S a$ead o> ti (e

8lassi# 6S VMGs

- Boal*
  - Run t$e 5inar) uest 6S +it$ lo+er $ard+are -rivile es t$an su-ervisor (ode3 itGs 5een desi ned >orF
- Means*
  - &rans-arent Virtuali%ation* *!>ull or native"*
    - 1o (odi>i#ation o> t$e 6S i (a e
    - Full) Virtuali%a5le Pro#essors !V&-'3 AMD-V3 I/M PP8"
    - D)na (i# /inar) &ranslation !VM+are"
  - Para-virtuali%ation*
    - Modi>i#ation o> so (e o> t$e 6S . A= sour#e >iles
    - *!#an 5e seen as a -ort to a ne+ -ro#essor ver) si (ilar to t$e real one"F*

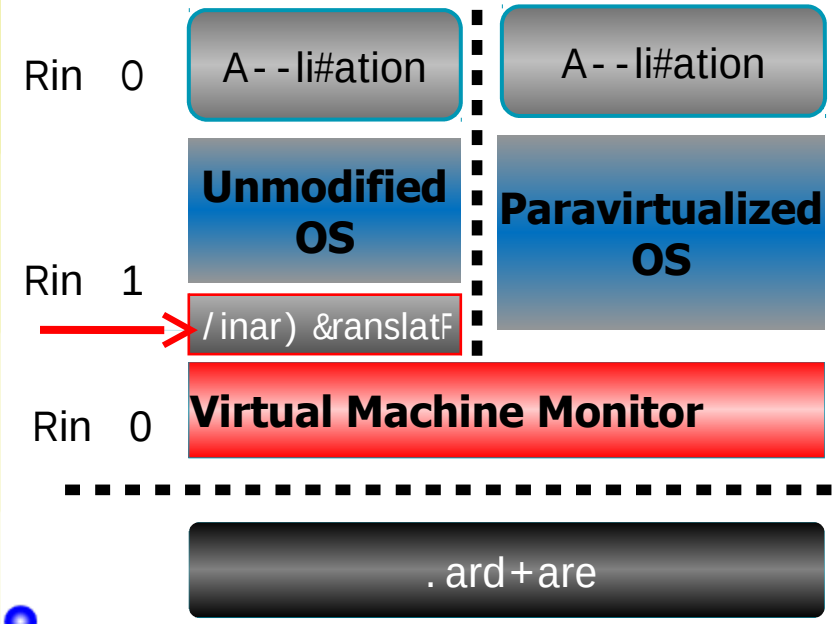
Full) Virtuali%a5le Pro#essors

- Issue*
  - Is t$ere an) instru#tions +$ose 5e$avior di>>ers or s$ould di>>er in t$ese 2 (odes,
- Sensitive Instru#tions*
  - &ra- to su-ervisor (ode +$en e'e#uted >ro (user (ode * 62
    - 0'* #li3 sti !Intel 'VK" tra- +$en e'e#uted >ro (user- (ode
  - 1o-o-* not 62*
    - 0'* P6PF !Intel 'VK"* Interru-t-ena5le >la re (ains una>>e#ted in user (ode
  - Bet s)ste (/ $ard+are status* not 62
    - 0'* Read 8RO !Intel 'VK" +ould return true -$)si#al in>o3 instead o> virtuali%ed in>oF

. ard+are assisted virtuali%ation

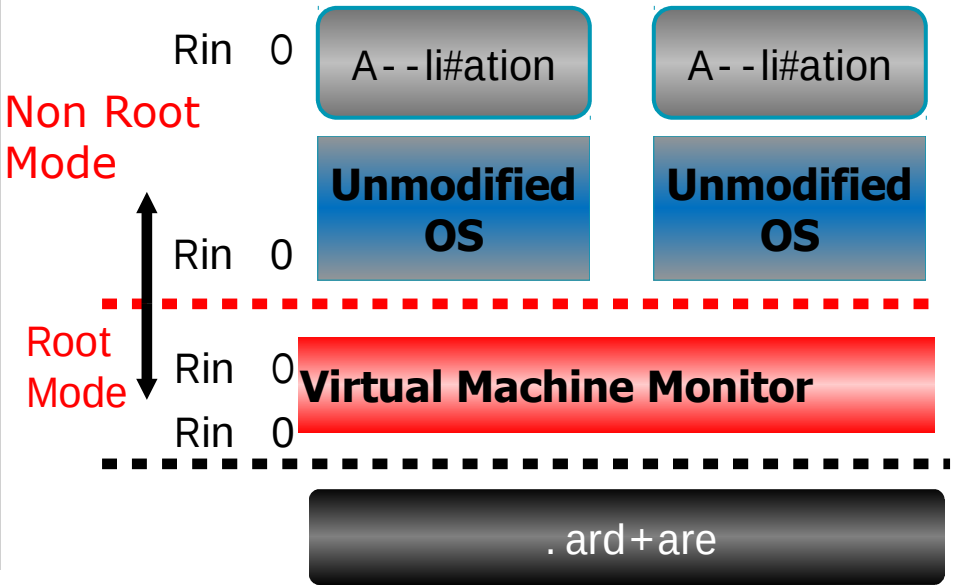
Non Hardware assisted

- /inar) translation
- Para-virtuali%ation
- 8o (-le'3 Sour#e #odeA



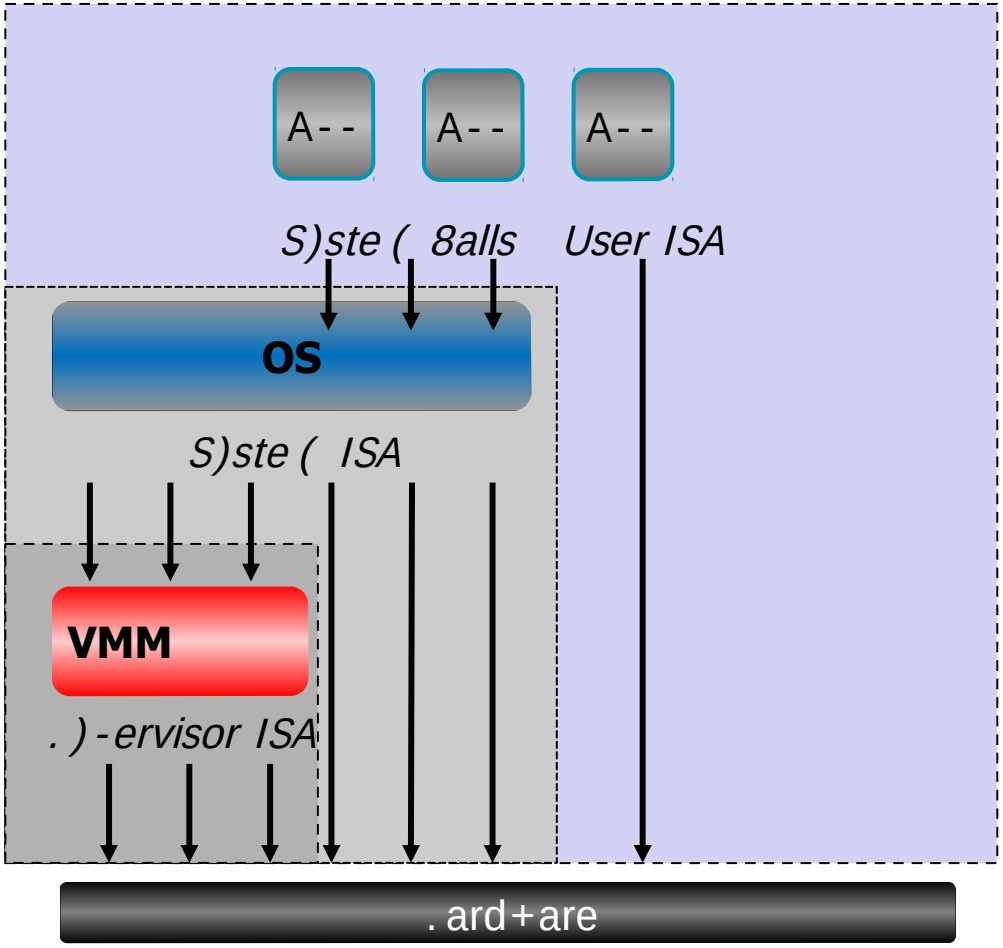
Virtualization Technology

- 2 1e+ e'e#ution (odes3
- And a set o> $ard+are-5ased tri ers to s+it#$ 5et+een t$e (



. ard+are assisted virtuali%ation

- &$e 6S $as not a##ess to t$e >ull ISA +$en runnin in a VMF



D)na (i# /inar) &ranslation

- 1on .ard+are assisted virtuali%ation* !'VK3A"
- Run un (odi›ied 5inar) uest in a less--rivile ed (ode !e'* rin 0 instead o› rin 0 on 'VK"
- D)na (i# 5inar) translation !eF F* VM+are on 'VK"
  - VMM d)na (i#all) 9re-+rites: -rivile ed instru#tions +$i#$ +ould 5e silentl) e'e#uted in user (odeF
  - 6n de (and3 #a#$ed
  - Me (or) #onsu (-tion i (-a#t
  - &i (in deter (inis (i (-a#t
- &rans-arent 5ut #o (-le' solution
  - #o (-le'it) resides in t$e VMM

Paravirtuali%ation

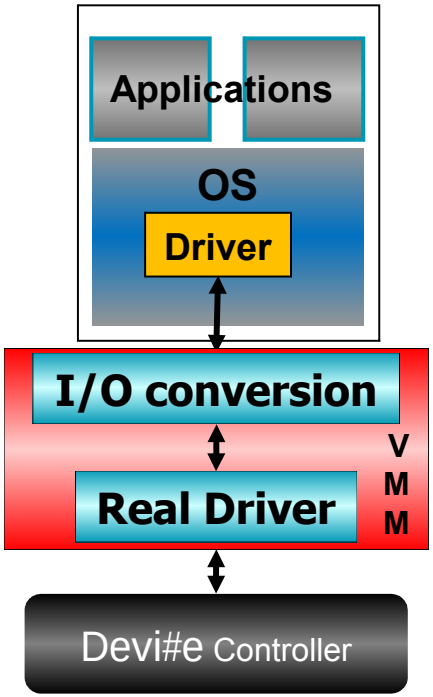
- Modi>) t$e .A= o> t$e uest 6S
- For se#ure isolation -ur-oses*
  - Rel) on (e(or) -rote#tion -rovided 5) $ard+are
  - Run uest 6S in less -rivile ed (ode
    - 0'* rin 1 instead o> rin 0 on 'VK
    - 0'* user (ode instead o> su-ervisor (ode on ARM
- /etter -er>or (an#e level t$an D)na(i# /inar) &ranslation!,'3 5ut intrusiveF
  - Solution o> #$oi#e +$en no $ard+are su--ort !PP83 ARM" and 6S (odi>i#ations are -ossi5le

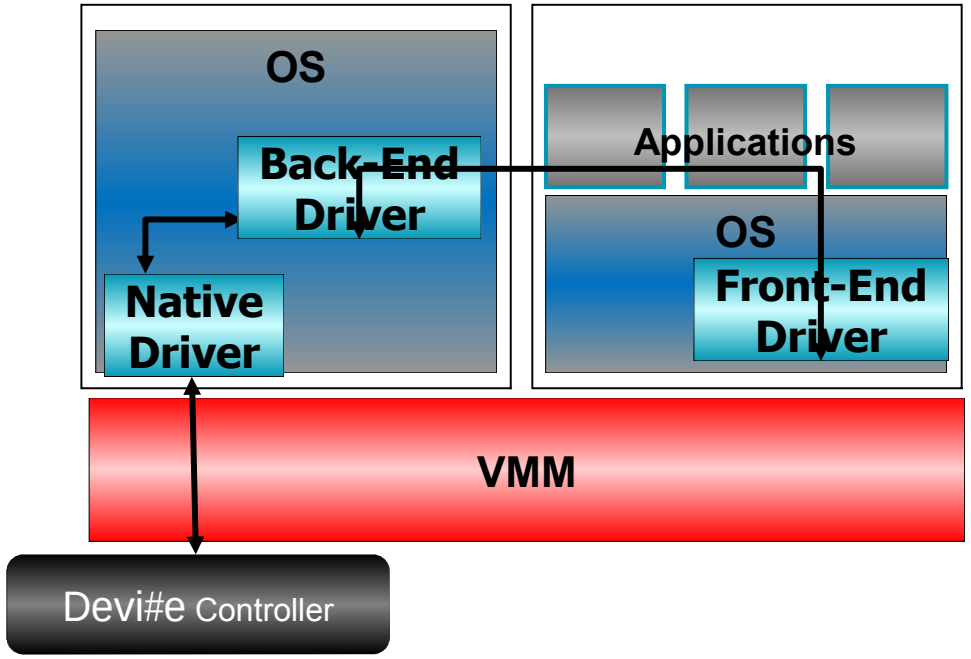
Virtuali%ation and devi#es

- 9S$ared: devi#es*
  - A##essed 5) (ore t$an one VM
  - 0'* dis7 is s$ared3 -artitions are not
  - 0'* 0t$ernet J a#tuall) 5rid in /routin 5et+een virtual and -$)si#al
- 1on s$ared devi#es
  - Devi#es used e'#lusivel) 5) a sin le VM
  - 0'* 1et+or7 inter>a#e
- Virtuali%ed 5) VMM
- Virtuali%ed +it$in a dedi#ated VM
  - Do (03 Do (I/6 in 4en3 9An): VM in V=4
- Dire#t -$)si#al devi#e a##ess >ro (VM
  - V&-d3 P8I su--ort / e'tensions 3 VMD@3 / V=43 A

Virtuali%ation and Devi#es

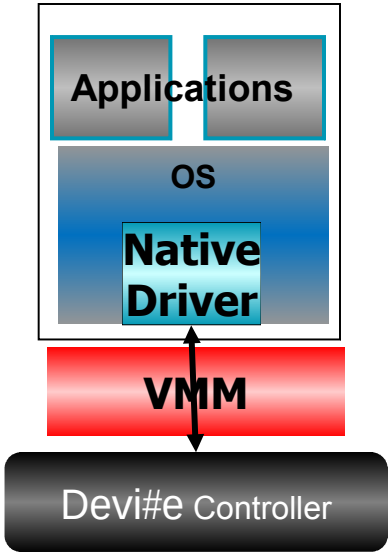
- Di>>erent +a)s to -rovide a##ess to devi#es*
- &rans-arent I/6Gs or 9-ara-virtuali%ed: I/6Gs
  - ProGs and 8onGs in 5ot$ #ases





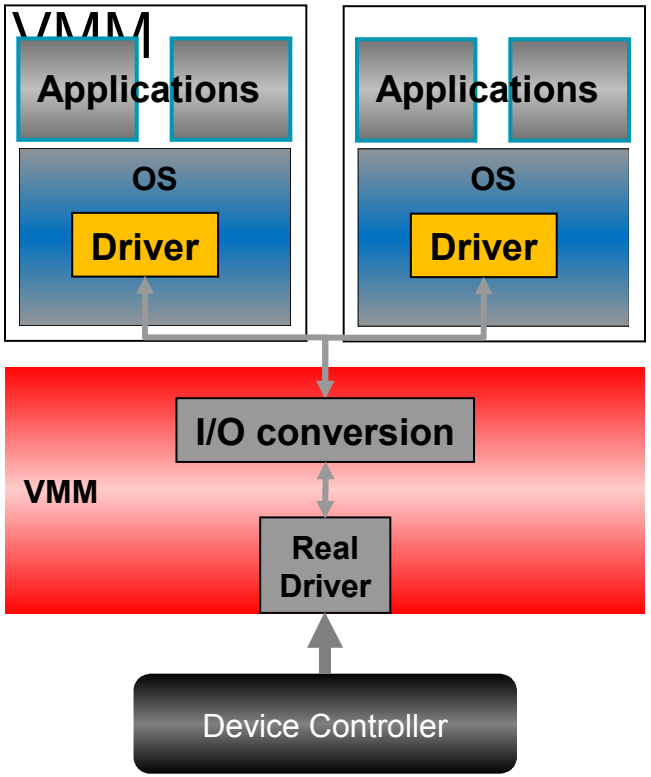
Virtuali%ation and Devi#es !8ontGd"

- /etter $ard+are su--ort*
  - P8I SRI6V3 MRI6V3
  - Intel V&-d3
  - S-e#i>i# #ontrollers !eF F* VMD@"
- 6r S-e#i>i# VMM i (-le (entations
  - V=4
  - SU1 I/6 =D6M !li (ited nu (5er"
- Un (odi>ied drivers3 5etter -er>or (an#e



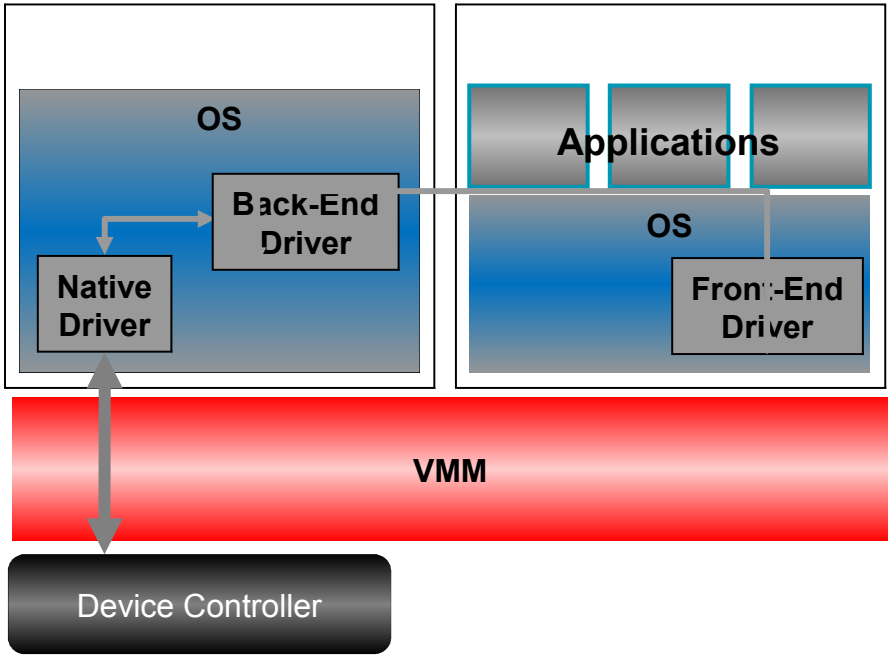
S$arin Devi#es

- S$ared devi#es are a #on#ern >or >ailure resilien#e
- S$ared devi#es -rovided 5) VMM*
  - Failure o> driver i(-lies >ailure o> VMM
  - And >ailures o> all VMGs



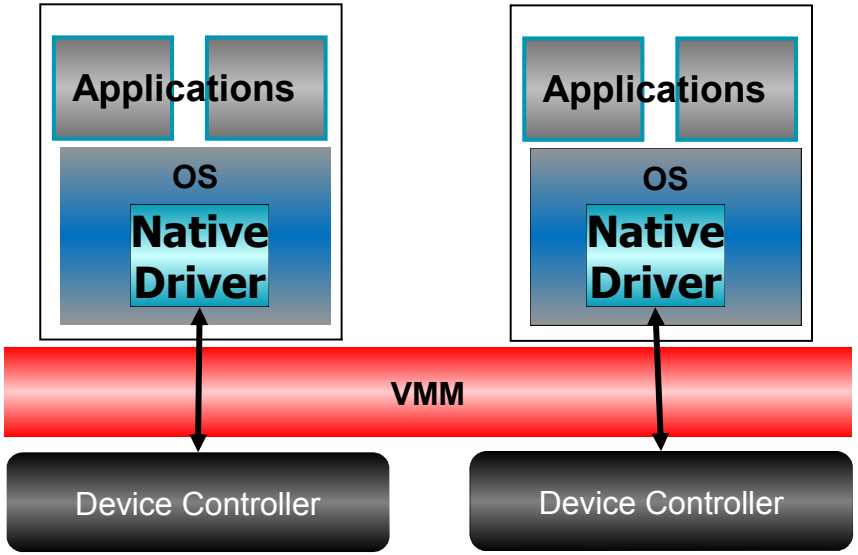
S$arin Devi#es

- S$arin -rovided 5) a VM3 t$rou $ 5a#7-end driver
- Failure o› driver PU ›ailure o› VM
- 6nl) #lient VMGs are i(-a#ted
- Restart under #ondition



1ot S$arin Devi#es

- Multi-le I/6 a5le VM6s #ould solve t$e de-enda5ilit) issue
- At t$e #ost o› (ore devi#es



A enda !1/2"

- Use #ases
- Multi-#ore #$allen es
- Virtuali%ation &a'ono ( )
  - S)ste (level3 Pro#ess level virtuali%ation3 6S virtuali%ation
  - 1ative versus . osted Virtuali%ation
  - &rans-arent virtuali%ation versus -ara-virtuali%ation
- 0 (5edded E Real-&i (e Virtuali%ation Re?uire (ents
- . ard+are 0volution

. ard+are 0volution

- 4VK
  - Multi-#ore3
  - . ard+are assisted virtuali%ation*
    - vt-'3 vt-i3 sv (3
    - Pa e ta5le su- -ort
    - I/6%s* vt-d3 io ((u3 !trans-arent virtuali%ed DMA"
    - H P8I* SRI6V3 MRI6VI3 !VMD@"
- Po+er Ar#$ite#ture >or 0 (5edded !@orI@"
  - Multi#ore -ro#essors
  - 0 rin s* user3 su-ervisor3 $)-ervisor
  - Address s-a#es ta ed +it$ ID
  - I/6 su- -ort via PAMU !si (ilar to vt-d / io ((u"
- Po+er >or Servers

. ard+are 0volution

- ARM
  - Road (a- to (ulti-#ore
  - 1o #urrent su--ort ›or virtuali%ation
  - . o+ever3 &rustRone -er (its to isolate 2 environ (ents
- SPAR8
  - 1ia ara* (ulti-#ore / (ulti-t$readed
  - . ard+are/Fir (+are su--ort ›or -ara-virtuali%ation

A enda !2/2"

- Des7to-3 Data 8enter solutions

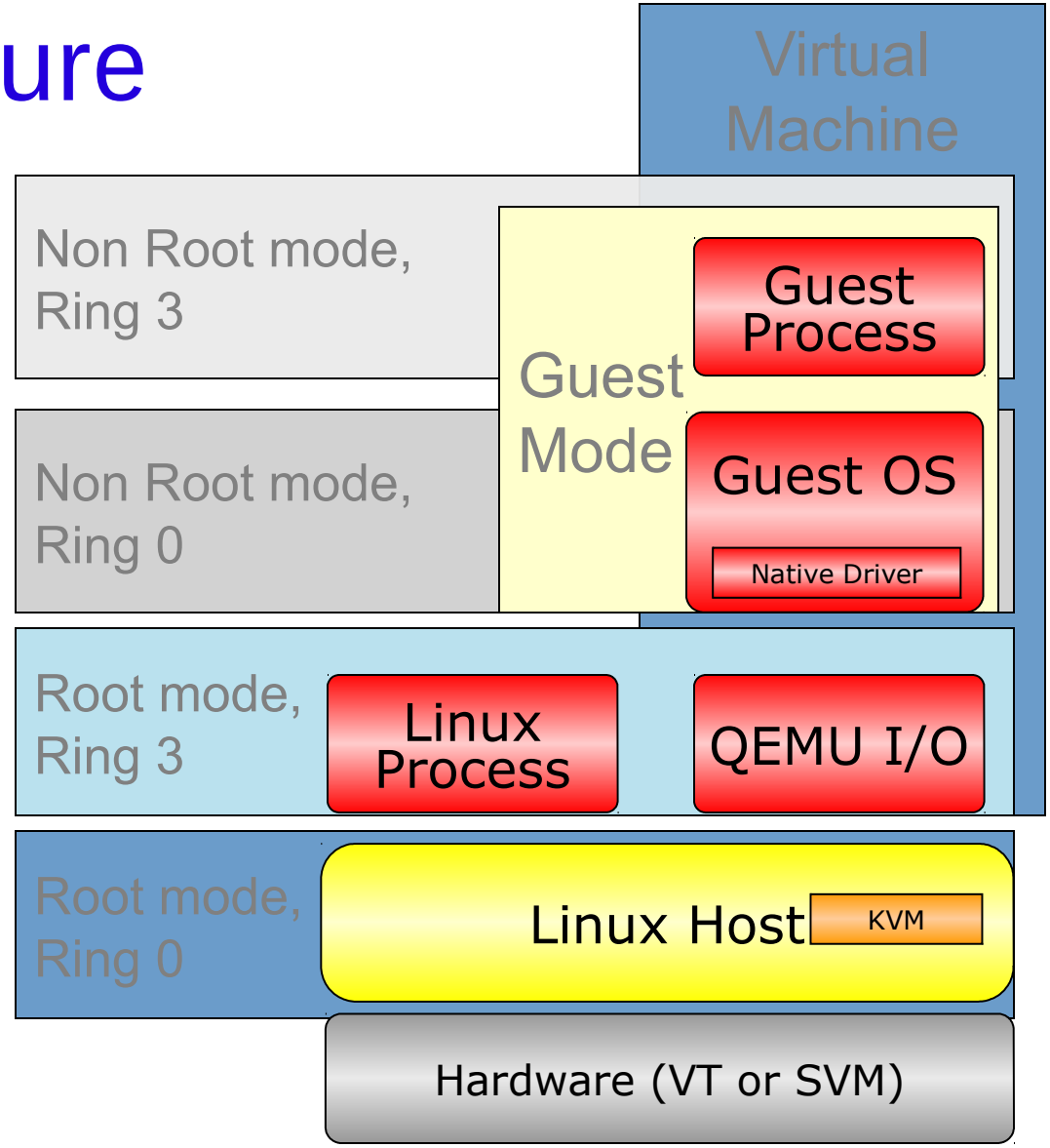
- Surve) o› 0 (5edded solutions*
  - .)5rid 6S6s3 8o-runnin 6S6s
  - As) ((etri# Multi-Pro#essin
  - Mi#ro-7ernel 5ased solutions
  - 0 (5edded Real-&i (e .) - ervisors
- Mis#ellaneous
  - &ools
  - . i $-Availa5ilit)
  - Standardi%ation

. osted Virtuali%ation*

- Start a $ost 6S >irst
- D)na (i#all) e'tends t$e 6S +it$ 7ernel (odule and a--li#ations -rovidin virtuali%ation
- Virtuali%ation relies on . ost 6S servi#es
- &)-i#all) 5inar) Buest 6S su--orted !&rans-arent Virtuali%ation"
- 1o isolation 5et+een . ost 6S and VMM
- 0'a (-les*
  - VM+are <S3 SVista3 Parallels3 Virtual/o'3 .)-er-V3 2VM
  - HUser-Mode =inu'I

2VM Ar#$ite#ture

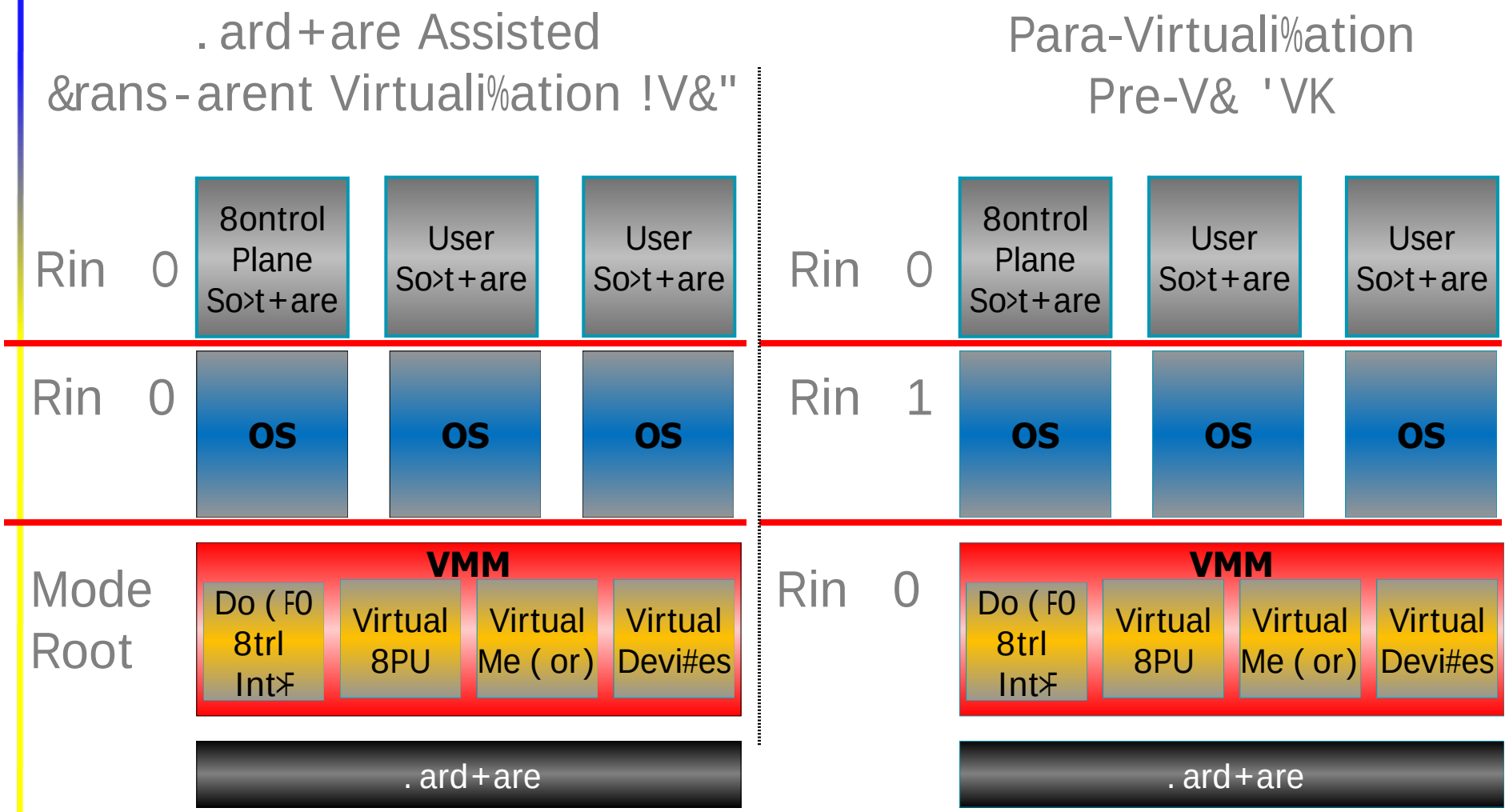
- Buest I/6 are tra--ed 5) 2VM and redire#ted to @0MUF
- VMGs are run and s#$eduled as =inu' .ost -ro#essesF



1ative Virtuali%ation

- Start a $)-ervisor / VMM on 5are (etal3
- &$en3 t)-i#all) start a 9#ontrol do(ain: !4en" or a Servi#e 8onsole !VM+are"
- 8an t$en Jd)na(i#all)- #reate VM6s
- Devi#e Virtuali%ation (a) 5e -rovided
  - /) VMM !VM+are 0S4"
  - /) ot$er VM6s !4en"
- 0'a(-les*
  - VM+are 0S4!i"3 4en3
  - V=4

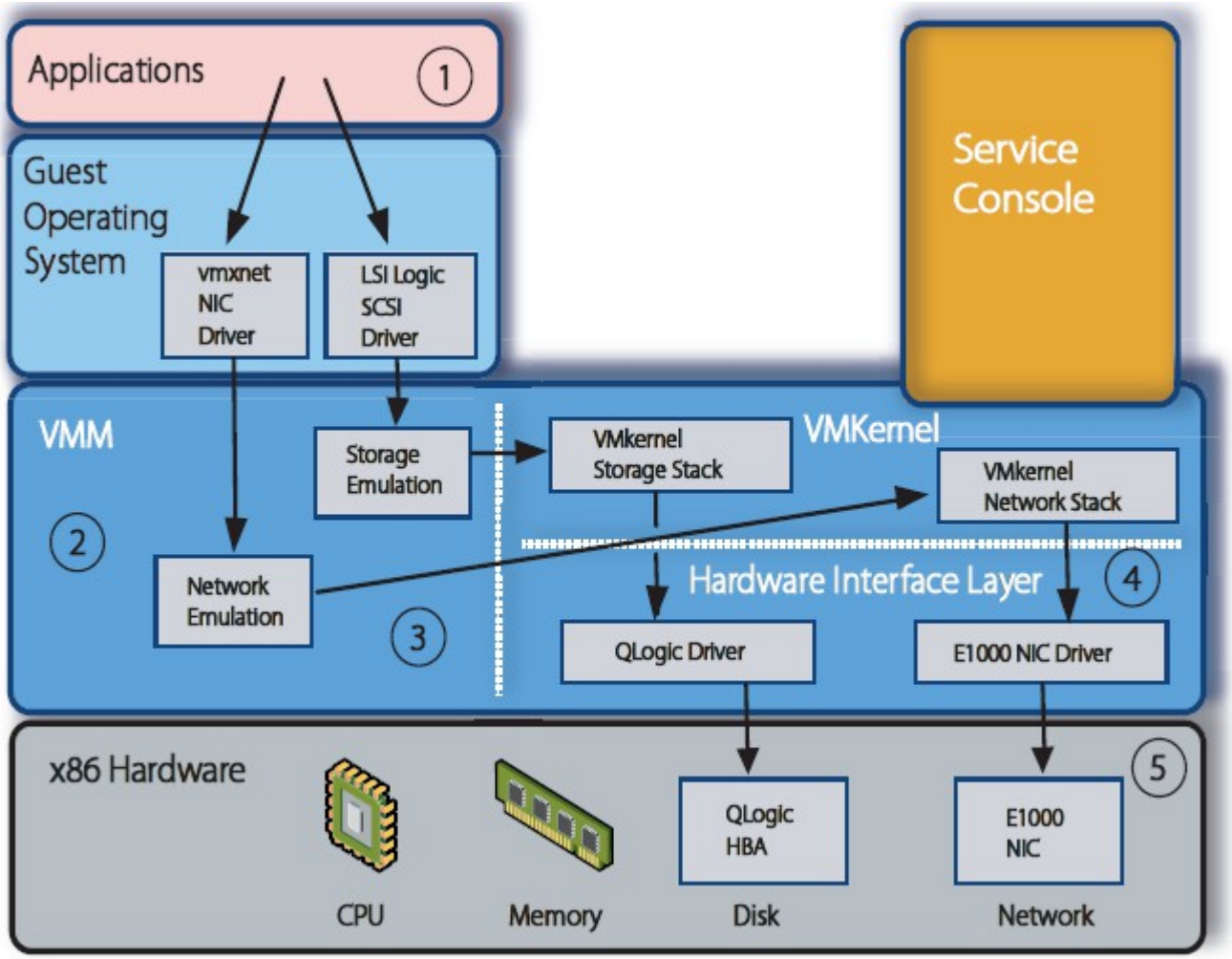
/ are Metal .) - ervisors* 401



401*

- Mostl) dedi#ated to Data 8enters* 'VK3 Itaniu (and I/M PP8
  - Paravirtuali%ation or . < assisted virtuali%ation eli (inates t$e need >or d)na (i# translation
  - Su--orts =inu'3 Free/SD3 PlanN3 A
  - Su--orts <indo+s on Intel V&3 AMD SVMF
  - &ools >or data #enter (ana e (ent ! (i ration3 -rovisionin FF"
- =i (ited su--ort >or ot$er -ro#essors
  - ARM in -ro ress
- 1o su--ort >or e (5edded #on>i urations
  - All VM%s #reated >ro (Do (0
  - 1eed to +ait >or #o (-lete initiali%ation o> Do (0

D)na (i# /inar) &ranslation* VM+are 0S4



/are-Metal .)-ervisors* 0S4

- D)na (i# /inar) &ranslation !6S un (odi›ied in rin 0"
  - So ›ar3 even on !02 5its" Intel V&-i3 AMD SVM
  - .ard+are virtuali%ation used on K; 5its (a#$ines
- 9/i :*
  - In#ludes Virtual Me (or) Mana e (ent3 File s)ste (!VMFS"3 1et+or7 sta#73 driversA
- Me (or) Mana e (ent*
  - P$)si#al Me (or) (a) 5e over #o ((itted !-a in out uest 6S%s"
  - D)na (i# #ontent 5ased -a e s$arin 3 Me (or) 5alloonin
- Dedi#ated to Data 8enters !'VK / Itaniu (onl)"

1ative versus .osted,

- VMM -rovides*
  - Me(or) (ana e(ent !even s+a- >or 0S4"3
  - S#$edulin !o> virtual 8PU6s on -$)si#al ones"
  - I/6 a##ess !virtual devi#es3 net+or7 -roto#ols3A"
  - Inter-VM #o((uni#ations
  - VM (ana e(ent* #reate3 sto-3 resu(e3 delete3A
  - Resour#e #ontrol3 a##ountin 3AF
- @uestion*
  - Isn6t t$at si(ilar to 6S servi#es >or -ro#esses, !W"
- @uestion*
  - <$at a5out a XVMM/VMM +ar3 to re-la) t$e 7ernel/X7ernel one, ☺

!W" Red .at initiall) ado-ted 4en3 is no+ usin 2VMF

=ive Mi ration o> VM

- Mi te a run nin VM >ro (one -$)si#al server to a t$er one +it$out loss o> servi#e ! (ini (al do+ ti (e * *=oad / a an#in 3 Preventive (aintenan#e*
  - ar: o> $ard+are #o (-ati3lit) #onstraints*

Virtuali%ation and #lusters

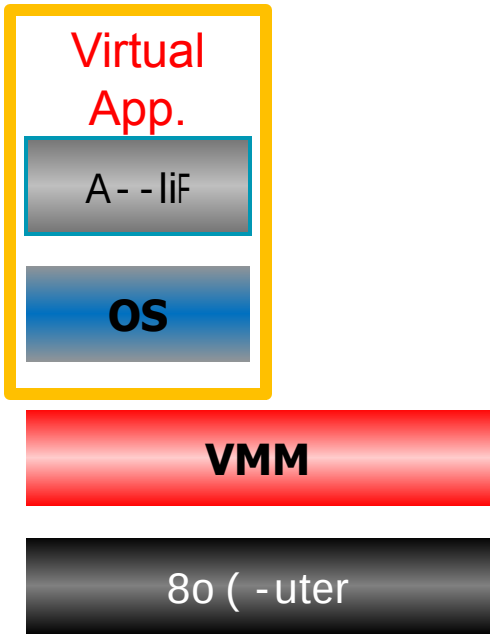
- 8entral (ana e (ent tools
  - 0' * VM+are virtual 8enter3 si (ilar tools >or ot$ers
- Resour#e #ontrols on VMGs
  - .ierar#$i#al resour#e -ools ! (e (or)3 8PU3A"
    - Ma) s-an -$)si#al (a#$ine 5oundaries
  - Si (ilar to resour#e (ana e (ent >or -ro#ess
- Auto (ati# restart o> >ailed VMs A so (e+$ere
  - 9.i $-Availa5ilit): solutions
- =ive Mi ration
- 9Sna-s$ot li5raries:

Virtuli%ation leads to Virtual A--lian#e

- Deliver A--li#ation +it$ its #ontainin Virtual Ma#$ine
  - 1o 6S version issue an) (ore
  - 1o #o (-ati5ilit) o› li5raries3 tools +it$ ot$er a--li#ations
  - A--li#ation alread) installed and -artl) #on›i ured
  - VM+are $as a lar e dire#tor) o› Virtual A--lian#es ›or VMPla)er
  - Bet rid o› -$)si#al devi#es de-enden#ies3 deal onl) +it$ virtual devi#es

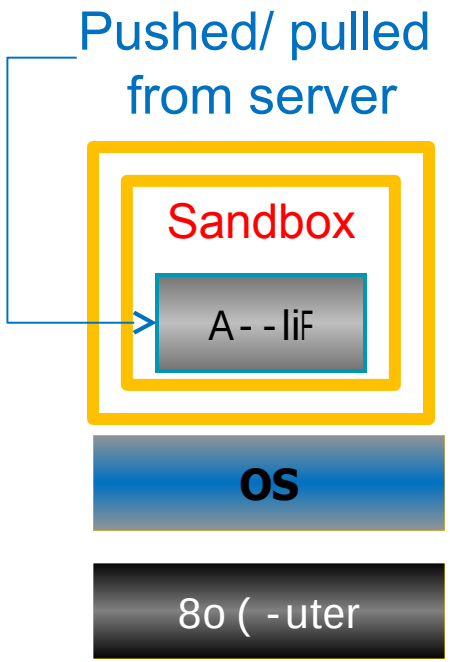
1e+ +a)s to loo7 at a- -li#ations

Virtual A- -lian#es

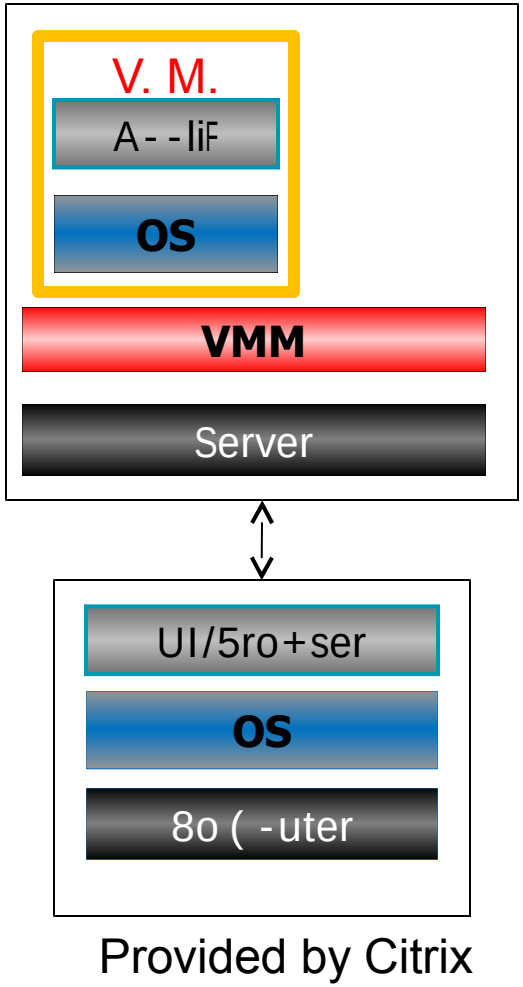


Promoted by VMware

Virtuali%ed A- -li#ations



Provided by MSFT/Softgrid



Provided by Citrix

0volution o> Virtual A- -lian#es

- A##ess to devi#es* virtuali%ed
  - 0it$er t$rou $ dedi#ated VM!4en"3 or 5) VMM!0S4"