Sun microsystems

Solaris Volume Manager : Userland Model

Commands

- Setup I18N stuff (locale, TEXT_DOMAIN)
- Init Cluster lib (sdssc_bind_library)
- md_init (opens /dev/md/admin and sets up to catch signals)
- Check for root privileges
- Parse input arguments
- Command specific code – usually calls into libmeta

Libraries

- libpreen
  - SVM library to provide a physical disk to fsck for optimal performance
  - Very fragile. Change at your own risk!!!
- libsvm
  - SVM library for install
- libmeta
  - SVM library for everything else

Libmeta breakdown

meta_check.c – checks components for other uses in the system (swap etc)
meta_db.c
meta_db_balance.c
Meta db code
meta_devadm.c – metadevadm specific code
meta_devstamp.c - timestamp
meta_error.c – error code
meta_getdevs.c – device specific
meta_hotspares.c – hotspare specific
meta_import.c
meta_statconcise.c
metaimport specific
meta_init.c – init of metadevices
meta_lib_prv.c – mnttab
meta_mem.c – memory alloca functions
meta_mirror.c
meta_mirror_resync.c
Mirror code
meta_mn_changelog.c
meta_mn_comm.d
meta_mn_handlers.c
meta_mn_msg_table.c
meta_mn_subr.c
Multi-node specific code (Oban)
meta_mount.c – mount point
meta_name.c – name manipulation
meta_nameinfo.c – get information given the name
meta_namespace.c – namespace utilities
meta_notify.c
meta_se_notify.c
Event notification

Libmeta breakdown (cont)

meta_patch.c patch /etc/vfstab & md.conf
meta_patch_root.c
meta_print.c – printing functions
meta_raid.c raid specific code
meta_raid_resync.c
meta_resync.c – raid and mirror resyncs
meta_rename.c – code for metarename
meta_replace.c – replace components in a metadevice
meta_reset.c – clear metadevices
meta_runtime.c – get runtime params from runtime.cf
meta_setup.c – setup facilities
meta_smf.c – smf interface
meta_sp.c – softpartition specific code
meta_stat.c – caching stat functions
meta_stripe.c – stripe specific code
meta_systemfile.c – interface with md.conf and mddb.cf
meta_tab.c – interface with md.tab
meta_time.c – timeof day
meta_trans.c – trans specific code
meta_userflags.c – get/set userflags for metadevices
metagetroot.c – get root device
metasplitname.c – split & splice device name
sdssc_bind.c – cluster binding

Libmeta breakdown (cont)

meta_mdcf.c – md.cf update

meta_set.c - generic
meta_set_drv.c - drive
meta_set_hst.c - host
meta_set_med.c - mediator
meta_set_prv.c -
meta_set_tkr.c – take and release

Set specific code

RPC functionality:

metarpcopen.c

metad_svc_stubs.c
meta_metad.c
meta_metad_subr.c

rpc.metad specific

meta_mh.c – rpc.metamhd

meta_med.c
meta_med_err.c

rpc.metamedd specific

meta_mh.c - rpc.metamhd

mdsetname_t

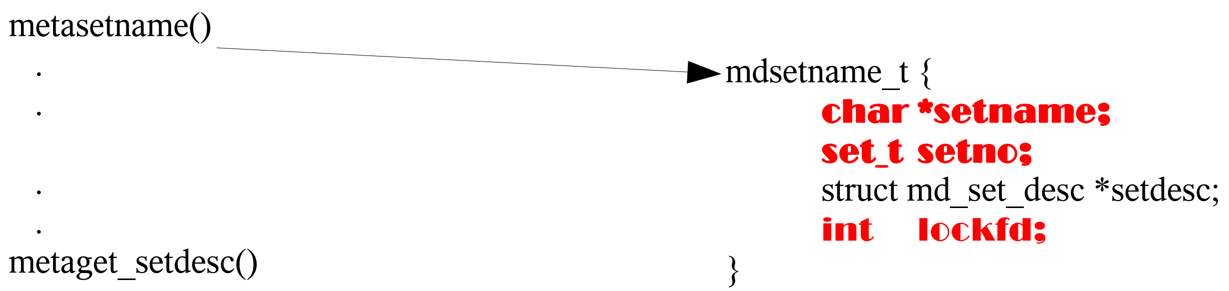
- Usually referenced as sp
- sp Obtained via call to metasetname(setname...)
  - Fills in setno, setname, lockfd
  - Does NOT fill in setdesc (yet)
- setdesc populated via metaget_setdesc
  - sp->setdesc != NULL , return it
  - sp->setdesc == NULL, call sr2setdesc
    - Uses daemon to get info from cached list (setrecords)

mdsetname_t (cont)

- set_snarf creates setrecords cached list
    - get_ur_record calls kernel for set record information
    - Set record information (sr) is added to the cached list

mdsetname_t

Fill in the setname, setno, and lockfd elements



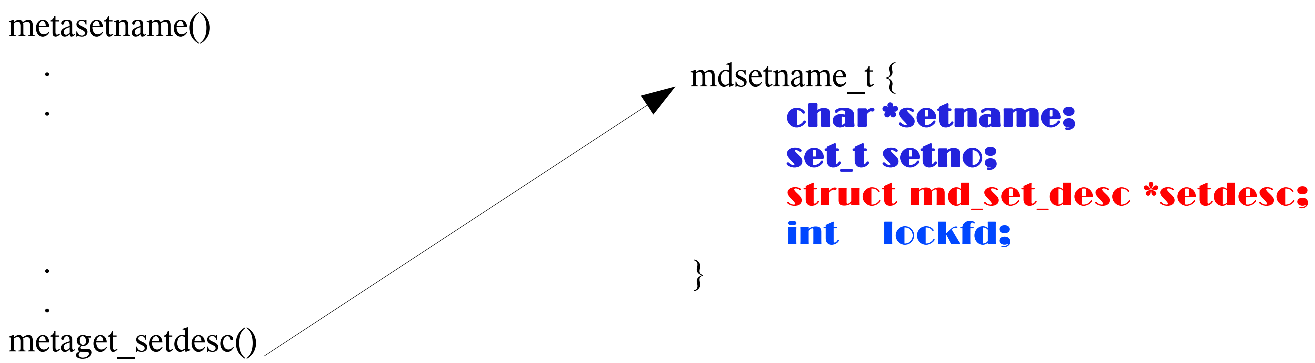
metasetname()
.
.
.
.
metaget_setdesc()

```
mdsetname_t {
        char *setname;
        set_t setno;
        struct md_set_desc *setdesc;
        int    lockfd;
}
```

Red-bold = field being filled in
Blue-bold = field previously filled in

mdsetname_t

Fill in the setdesc pointer



```
metasetname()
    .
    .


    .
    .
metaget_setdesc()
```

```
mdsetname_t {
        char *setname;
        set_t setno;
        struct md_set_desc *setdesc;
        int    lockfd;
}
```

mdname_t

- Usually referenced as np
- “fast” functions
  - Use when DKIOCGGEOM info isn't needed
- Use metainitname
  - Memory must be allocated before call
  - 0's all fields except dev, key, start_blk and end_blk (default values)

mdname_t – init of stripe

Allocate and initialize the mdname_t structure

```
meta_init_stripe()
    metaname()
        metaname_common()
            meta_name_getname()
            getrawnames()
            metainitdrivename()
            getparts()
            metainitname()
            setup_slice()
                getnames()
    metachkmeta()
    metagetmiscname()
        meta_getminor()
        metaioctl(MD_IOCGET_DRVNM)
    parse command line options
    build entries for each comp in memory
    meta_create_stripe()
        walk each component()
            metagetsize()
                metagetvtoc()
            metagetstart()
            add_key_name()
```

```
mdname_t{
        struct mddrivename_t        *drivenamep;
        char    *cname;
        char    *bname;
        char    *rname;
        char    *devicesname;
        char    *minor_name;
        md_key64_t  dev;
        mdkey_t     key;
        diskaddr_t  end_blk;
        diskaddr_t  start_blk;
}
```

mdname_t – init of stripe

Fill in the drivenamep field

```
meta_init_stripe()
    metaname()
        metaname_common()
            meta_name_getname()
            getrawnames()
            metainitdrivename()
            getparts()
            metainitname()
            setup_slice()
                getnames()
    metachkmeta()
    metagetmiscname()
        meta_getminor()
        metaioctl(MD_IOCGET_DRVNM)
    parse command line options
    build entries for each comp in memory
    meta_create_stripe()
        walk each component()
            metagetsize()
                metagetvtoc()
            metagetstart()
            add_key_name()
```

```
mdname_t{
        struct mddrivename_t     *drivenamep;
        char     *cname;
        char     *bname;
        char     *rname;
        char     *devicesname;
        char     *minor_name;
        md_key64_t  dev;
        mdkey_t     key;
        diskaddr_t  end_blk;
        diskaddr_t  start_blk;
}
```

mdname_t – init of stripe

Fill in cname, bname, rname and dev elements

```
meta_init_stripe()
    metaname()
        metaname_common()
            meta_name_getname()
            getrawnames()
            metainitdrivename()
            getparts()
            metainitname()
            setup_slice()
                getnames()
    metachkmeta()
    metagetmiscname()
        meta_getminor()
        metaioctl(MD_IOCGET_DRVNM)
    parse command line options
    build entries for each comp in memory
    meta_create_stripe()
        walk each component()
            metagetsize()
                metagetvtoc()
            metagetstart()
            add_key_name()
```

```
mdname_t{
        struct mddrivename_t     *drivenamep;
        char *cname;
        char *bname;
        char *rname;
        char    *devicesname;
        char    *minor_name;
        md_dev64_t      dev;
        mdkey_t         key;
        diskaddr_t      end_blk;
        diskaddr_t      start_blk;
}
```

mdname_t – stripe via metastat

Fill in drivenamep, cname, bname, rname elements

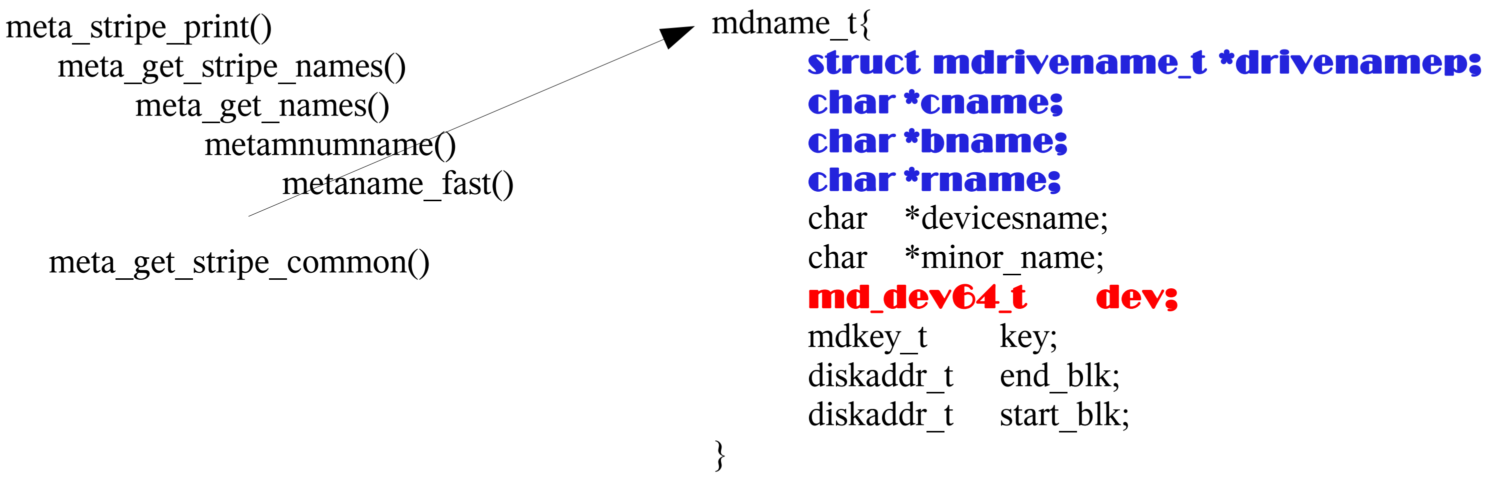
```
meta_stripe_print()
    meta_get_stripe_names()
          meta_get_names()
               metamnumname()
                     metaname_fast()
                         metaname_common()
                             metainitfastname()

    meta_get_stripe_common()
```

```
mdname_t{
        struct mdrivename_t *drivenamep;
        char *cname;
        char *bname;
        char *rname;
        char    *devicesname;
        char    *minor_name;
        md_dev64_t  dev;
        mdkey_t      key;
        diskaddr_t   end_blk;
        diskaddr_t   start_blk;
}
```

mdname_t – stripe via metastat

Fill in dev element



```
meta_stripe_print()
      meta_get_stripe_names()
               meta_get_names()
                       metamnumname()
                                metaname_fast()

     meta_get_stripe_common()
```

```
mdname_t{
           struct mdrivename_t *drivenamep;
           char *cname;
           char *bname;
           char *rname;
           char    *devicesname;
           char    *minor_name;
           md_dev64_t      dev;
           mdkey_t       key;
           diskaddr_t    end_blk;
           diskaddr_t    start_blk;
}
```

mddrivename_t

- Usually referred to as dnp
- Type field indicates if a metadevice or disk

mddrivename_t – init of stripe (metadevice)

Allocate and initialize the mddrivename_t structure

```
meta_init_stripe()
    metaname()
        metaname_common()
            meta_name_getname()
            getrawnames()
            metainitdrivename()
            getparts()
            metainitname()
    metachkmeta()
    metagetmiscname()
        meta_getminor()
        metaioctl(MD_IOCGET_DRVNM)
    parse command line options
    build entries for each comp in memory
    meta_create_stripe()
        walk each component()
            metagetsize()
                metagetvtoc()
            metagetstart()
            add_key_name()
```

```
mddrivename_t{
        char        *not_used;
        char        *cname;
        char        *rname;
        mdnmtype_t  type;
        char        *devid;
        int         errnum;
        mdgeom_t        geom;
        mdcinfo_t       cinfo;
        mdvtoc_t        vtoc;
        struct  {
                u_int   parts_len;
                mdname_t        *parts_val;
        } parts;
        mdsidename_t *side_names;
        mdkey_t         side_names_key;
        char            *miscname;
        struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (metadevice)

Fill in type element

```
meta_init_stripe()
        metaname
              metaname_common
                     meta_name_getname
                     getrawnames
                     metainitdrivename
                     getparts
                     metainitname
        metachkmeta
        metagetmiscname
             meta_getminor
             metaioctl(MD_IOCGET_DRVNM)
        parse command line options
        build entries for each comp in memory
        meta_create_stripe
              walk each component
                      metagetsize
                          metagetvtoc
                      metagetstart
                      add_key_name
```

```
mddrivename_t{
          char    *not_used;
          char    *cname;
          char    *rname;
          mdnmtype_t  type;
          char    *devid;
          int     errnum;
          mdgeom_t     geom;
          mdcinfo_t    cinfo;
          mdvtoc_t     vtoc;
          struct  {
                    u_int   parts_len;
                    mdname_t     *parts_val;
          } parts;
          mdsidename_t *side_names;
          mdkey_t      side_names_key;
          char         *miscname;
          struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (metadevice)

Fill in the parts structure element

```
meta_init_stripe()
        metaname
              metaname_common
                     meta_name_getname
                     getrawnames
                     metainitdrivename
                     getparts
                     metainitname
        metachkmeta
        metagetmiscname
             meta_getminor
             metaioctl(MD_IOCGET_DRVNM)
        parse command line options
        build entries for each comp in memory
        meta_create_stripe
              walk each component
                      metagetsize
                          metagetvtoc
                      metagetstart
                      add_key_name
```

```
mddrivename_t{
         char    *not_used;
         char    *cname;
         char    *rname;
         mdnmtype_t  type;
         char    *devid;
         int     errnum;
         mdgeom_t     geom;
         mdcinfo_t    cinfo;
         mdvtoc_t     vtoc;
         struct       {
                  u_int parts_len;
                  mdname_t        *parts_val;
         } parts;
         mdsidename_t *side_names;
         mdkey_t       side_names_key;
         char          *miscname;
         struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (metadevice)

Fill in the cname and rname elements

```
meta_init_stripe()
    metaname
        metaname_common
            meta_name_getname
            getrawnames
            metainitdrivename
            getparts
            metainitname

    metachkmeta
    metagetmiscname
        meta_getminor
        metaioctl(MD_IOCGET_DRVNM)
    parse command line options
    build entries for each comp in memory
    meta_create_stripe
        walk each component
            metagetsize
                metagetvtoc
            metagetstart
            add_key_name
```

```
mddrivename_t{
        char    *not_used;
        char *cname;
        char *rname;
        mdnmtype_t  type;
        char    *devid;
        int     errnum;
        mdgeom_t    geom;
        mdcinfo_t   cinfo;
        mdvtoc_t    vtoc;
        struct      {
                u_int parts_len;
                mdname_t        *parts_val;
        } parts;
        mdsidename_t *side_names;
        mdkey_t     side_names_key;
        char        *miscname;
        struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (physical dev)

Allocate and initialize the mddrivename_t structure

```
meta_init_stripe()
        metaname()
        metachkmeta()
        build entries for rows
             build entries for components
                    metaname()
                                meta_name_getname
                                getrawnames()
                                Zalloc()
                                metainitdrivename()
                                getparts()
                                setup_slice()
                    meta_create_stripe()
```

```
mddrivename_t{
            char    *not_used;
            char    *cname;
            char    *rname;
            mdnmtype_t  type;
            char    *devid;
            int      errnum;
            mdgeom_t      geom;
            mdcinfo_t     cinfo;
            mdvtoc_t      vtoc;
            struct  {
                    u_int   parts_len;
                    mdname_t      *parts_val;
            } parts;
            mdsidename_t *side_names;
            mdkey_t          side_names_key;
            char                 *miscname;
            struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (physical dev)

Fill in the type, devid, geom, and vtoc elements

```
meta_init_stripe()
        metaname()
        metachkmeta()
        build entries for rows
             build entries for components
                   metaname()
                              meta_name_getname
                              getrawnames()
                              Zalloc()
                              metainitdrivename()
                              getparts()
                              setup_slice()
                   meta_create_stripe()
```

```
mddrivename_t{
          char    *not_used;
          char    *cname;
          char    *rname;
          mdnmtype_t  type;
          char *devid;
          int       errnum;
          mdgeom_tgeom;
          mdcinfo_t       cinfo;
          mdvtoc_t  vtoc;
          struct  {
                    u_int   parts_len;
                    mdname_t      *parts_val;
          } parts;
          mdsidename_t *side_names;
          mdkey_t          side_names_key;
          char             *miscname;
          struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (physical dev)

Fill in the parts structure

```
meta_init_stripe()
        metaname()
        metachkmeta()
        build entries for rows
             build entries for components
                   metaname()
                         meta_name_getname
                         getrawnames()
                         Zalloc()
                         metainitdrivename()
                         getparts()

                         setup_slice()
                   meta_create_stripe()
```

```
mddrivename_t{
        char    *not_used;
        char    *cname;
        char    *rname;
        mdnmtype_t   type;
        char *devid;
        int       errnum;
        mdgeom_tgeom;
        mdcinfo_t      cinfo;
        mdvtoc_t  vtoc;
        struct        {
                u_int parts_len;
                mdname_t          *parts_val;
        } parts;
        mdsidename_t *side_names;
        mdkey_t        side_names_key;
        char           *miscname;
        struct md_common_t  *unitp;
}
```

mddrivename_t – init of stripe (physical dev)

Fill in the cname and rname elements

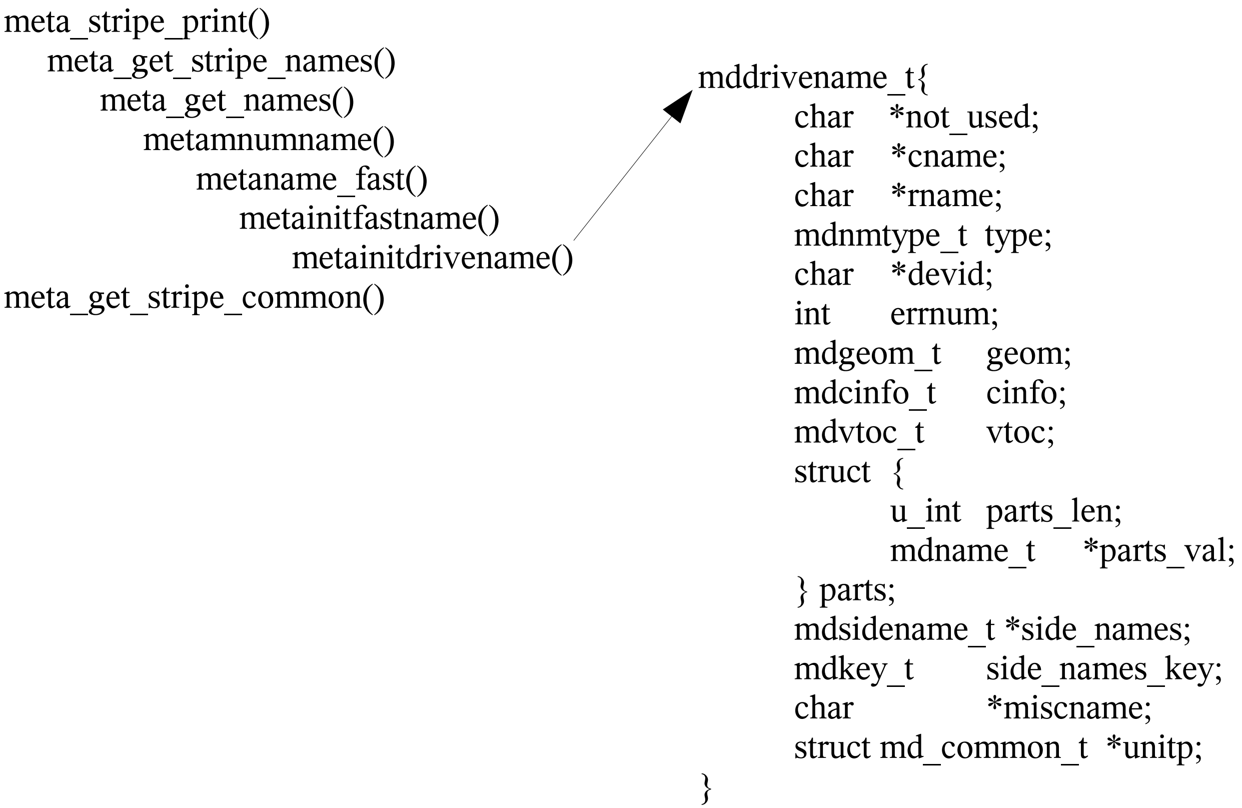
```
meta_init_stripe()
        metaname()
        metachkmeta()
        build entries for rows
             build entries for components
                   metaname()
                         meta_name_getname
                         getrawnames()
                         Zalloc()
                         metainitdrivename()
                         getparts()
                         setup_slice()

                   meta_create_stripe()
```

```
mddrivename_t{
            char    *not_used;
            char *cname;
            char *rname;
            mdnmtype_t  type;
            char *devid;
            int      errnum;
            mdgeom_tgeom;
            mdcinfo_t      cinfo;
            mdvtoc_t  vtoc;
            struct        {
                        u_int parts_len;
                        mdname_t          *parts_val;
            } parts;
            mdsidename_t *side_names;
            mdkey_t        side_names_key;
            char              *miscname;
            struct md_common_t  *unitp;
}
```

mddrivename_t – metastat (metadevice)

Allocate and initialize the mddrivename_t structure



```
meta_stripe_print()
    meta_get_stripe_names()
        meta_get_names()
            metamnumname()
                metaname_fast()
                    metainitfastname()
                        metainitdrivename()
meta_get_stripe_common()
```

```
mddrivename_t{
        char    *not_used;
        char    *cname;
        char    *rname;
        mdnmtype_t  type;
        char    *devid;
        int     errnum;
        mdgeom_t        geom;
        mdcinfo_t       cinfo;
        mdvtoc_t        vtoc;
        struct  {
                u_int   parts_len;
                mdname_t        *parts_val;
        } parts;
        mdsidename_t *side_names;
        mdkey_t         side_names_key;
        char            *miscname;
        struct md_common_t  *unitp;
}
```

mddrivename_t – metastat (metadevice)

Fill in the cname, rname, and type elements

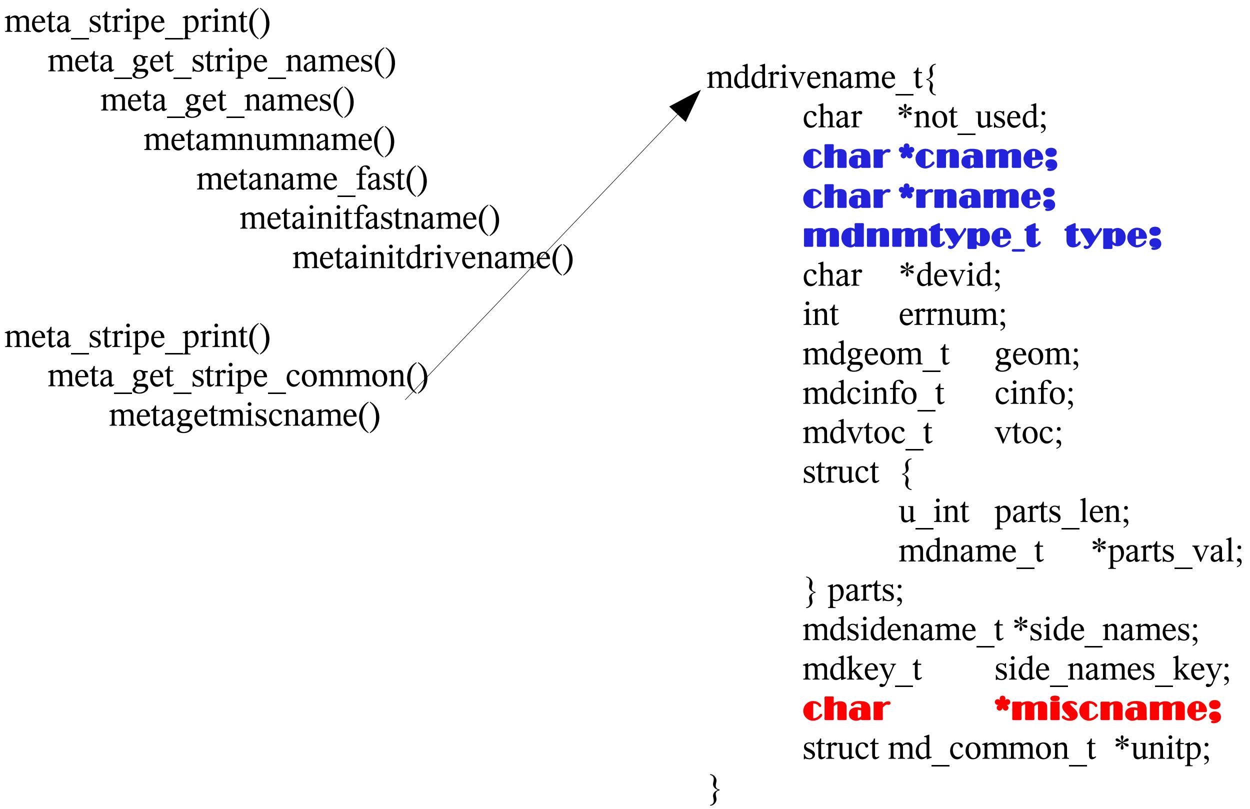
```
meta_stripe_print()
    meta_get_stripe_names()
         meta_get_names()
             metamnumname()
                  metaname_fast()
                      metainitfastname()
                           metainitdrivename()

meta_stripe_print()
    meta_get_stripe_common()
          metagetmiscname()
```

```
mddrivename_t{
        char    *not_used;
        char *cname;
        char *rname;
        mdnmtype_t  type;
        char    *devid;
        int     errnum;
        mdgeom_t    geom;
        mdcinfo_t   cinfo;
        mdvtoc_t    vtoc;
        struct  {
                u_int   parts_len;
                mdname_t     *parts_val;
        } parts;
        mdsidename_t *side_names;
        mdkey_t        side_names_key;
        char           *miscname;
        struct md_common_t  *unitp;
}
```

mddrivename_t – metastat (metadevice)

Fill in the miscname element



```
meta_stripe_print()
    meta_get_stripe_names()
         meta_get_names()
             metamnumname()
                  metaname_fast()
                      metainitfastname()
                           metainitdrivename()

meta_stripe_print()
    meta_get_stripe_common()
          metagetmiscname()
```

```
mddrivename_t{
         char    *not_used;
         char *cname;
         char *rname;
         mdnmtype_t  type;
         char    *devid;
         int     errnum;
         mdgeom_t    geom;
         mdcinfo_t   cinfo;
         mdvtoc_t    vtoc;
         struct  {
                 u_int  parts_len;
                 mdname_t     *parts_val;
         } parts;
         mdsidename_t *side_names;
         mdkey_t      side_names_key;
         char         *miscname;
         struct md_common_t  *unitp;
}
```

mddrivename_t – metastat (metadevice)

Fill in the unitp

```
meta_stripe_print()
    meta_get_stripe_names()
         meta_get_names()
             metamnumname()
                  metaname_fast()
                      metainitfastname()
                           metainitdrivename()

meta_stripe_print()
    meta_get_stripe_common()
          metagetmiscname()
          meta_get_mdunit()
         fill in stripe unit structure

     stripe_report()
```

```
mddrivename_t{
         char    *not_used;
         char *cname;
         char *rname;
         mdnmtype_t  type;
         char    *devid;
         int      errnum;
         mdgeom_t      geom;
         mdcinfo_t     cinfo;
         mdvtoc_t      vtoc;
         struct  {
                  u_int  parts_len;
                  mdname_t      *parts_val;
         } parts;
         mdsidename_t *side_names;
         mdkey_t         side_names_key;
         char            *miscname;
         struct md_common_t  *unitp;
}
```

mddrivename_t – metastat (physical dev)

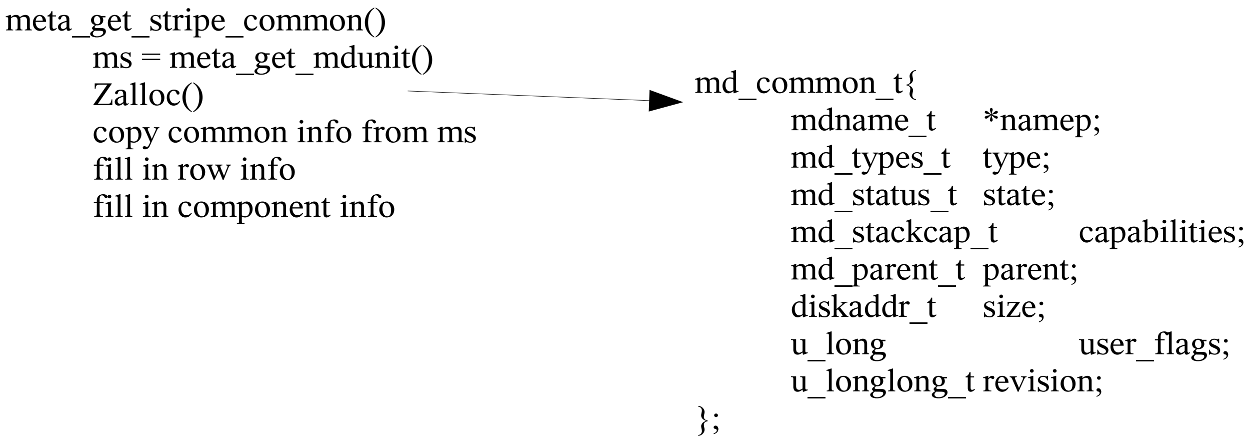
Fill in cname, rname, and type elements

```
meta_stripe_print()
    meta_get_stripe_names()
         meta_get_names()
             metamnumname()
                  metaname_fast()
                      metainitfastname()
                           metainitdrivename()
meta_stripe_print()
    meta_get_stripe_common()
          metagetmiscname()
          meta_get_mdunit()
         fill in stripe unit structure
         metakeyname()
             metaname_fast()
                  metaname_common()
                      metainitfastname()
     stripe_report()
```

```
mddrivename_t{
         char    *not_used;
         char *cname;
         char *rname;
         mdnmtype_t  type;
         char    *devid;
         int     errnum;
         mdgeom_t     geom;
         mdcinfo_t    cinfo;
         mdvtoc_t     vtoc;
         struct  {
                  u_int  parts_len;
                  mdname_t     *parts_val;
         } parts;
         mdsidename_t *side_names;
         mdkey_t        side_names_key;
         char           *miscname;
         struct md_common_t  *unitp;
}
```

md_common_t – stripe via metastat

Allocate and initialize the md_common_t structure



```
meta_get_stripe_common()
        ms = meta_get_mdunit()
        Zalloc()
        copy common info from ms
        fill in row info
        fill in component info
```

```
md_common_t{
        mdname_t        *namep;
        md_types_t      type;
        md_status_t     state;
        md_stackcap_t           capabilities;
        md_parent_t     parent;
        diskaddr_t      size;
        u_long                  user_flags;
        u_longlong_t revision;
};
```

md_common_t – stripe via metastat

Fill in the namep, type, state, capabilities, parent, size, user_flags and revision elements

```
meta_get_stripe_common()
        ms = meta_get_mdunit()
        Zalloc()
        copy common info from ms
        fill in row info
        fill in component info
```

```
md_common_t{
        mdname_t        *namep;
        md_types_t      type;
        md_status_t     state;
        md_stackcap_t   capabilities;
        md_parent_t     parent;
        diskaddr_t      size;
        u_long          user_flags;
        u_longlong_t    revision;
};
```

Error handling

- md_error_t usually referenced as ep
- Code is located in meta_error.c
- Many different “classes” of errors to accommodate different amount and types of info that needs to be passed along
- Each class has it's own processing function

Classes of errors

- sys – system error
- rpc – rpc error
- dev – device error
- use – usage error
- comp – component error
- hsp - hotspare pool error
- hs – hotspare error
- mddb – mddb error
- ds – diskset error
- overlap – overlap error

Error specific functions

- mdclrerror – clears the error structure
- md<class> error – fills in the class specific error information
- mdcookerror – fills in the basic error information
- mdstealerror – copy an error code from kernel to userland
- Various functions to print the message given the error code

SVM : Userland Model