NAME

domainset(9) - domainset functions and operation

SYNOPSIS

#include <sys/_domainset.h>
#include <sys/domainset.h>

```
struct domainset {
        domainset_t     ds_mask;
        uint16_t        ds_policy;
        domainid_t      ds_prefer;
                ...
};
```

*struct domainset **
DOMAINSET_FIXED(*domain*);

*struct domainset **
DOMAINSET_FT();

*struct domainset **
DOMAINSET_IL();

*struct domainset **
DOMAINSET_RR();

*struct domainset **
DOMAINSET_PREF(*domain*);

*struct domainset **
domainset_create(*const struct domainset *key*);

int
sysctl_handle_domainset(*SYSCTL_HANDLER_ARGS*);

DESCRIPTION

The **domainset(9)** API provides memory domain allocation policy for NUMA machines. Each *domainset* contains a bitmask of allowed domains, an integer policy, and an optional preferred domain. Together, these specify a search order for memory allocations as well as the ability to restrict threads and objects to a subset of available memory domains for system partitioning and resource management.

Every thread in the system and optionally every *vm_object_t*, which is used to represent files and other memory sources, has a reference to a *struct domainset*. The domainset associated with the object is consulted first and the system falls back to the thread policy if none exists.

The allocation policy has the following possible values:

DOMAINSET_POLICY_ROUNDROBIN
: Memory is allocated from each domain in the mask in a round-robin fashion. This distributes bandwidth evenly among available domains. This policy can specify a single domain for a fixed allocation.

DOMAINSET_POLICY_FIRSTTOUCH
: Memory is allocated from the node that it is first accessed on. Allocation falls back to round-robin if the current domain is not in the allowed set or is out of memory. This policy optimizes for locality but may give pessimal results if the memory is accessed from many CPUs that are not in the local domain.

DOMAINSET_POLICY_PREFER
: Memory is allocated from the node in the *prefer* member. The preferred node must be set in the allowed mask. If the preferred node is out of memory the allocation falls back to round-robin among allowed sets.

DOMAINSET_POLICY_INTERLEAVE
: Memory is allocated in a striped fashion with multiple pages allocated to each domain in the set according to the offset within the object. The strip width is object dependent and may be as large as a super-page (2MB on amd64). This gives good distribution among memory domains while keeping system efficiency higher and is preferential to round-robin for general use.

The **DOMAINSET_FIXED**(), **DOMAINSET_FT**(), **DOMAINSET_IL**(), **DOMAINSET_RR**() and **DOMAINSET_PREF**() macros provide pointers to global pre-defined policies for use when the desired policy is known at compile time. **DOMAINSET_FIXED**() is a policy which only permits allocations from the specified domain. **DOMAINSET_FT**() is a policy which attempts to allocate memory local to the current CPU, falling back to a round-robin policy if the initial allocation fails. **DOMAINSET_IL**() and **DOMAINSET_RR**() provide round-robin selection among all domains in the system, corresponding to the DOMAINSET_POLICY_INTERLEAVE and DOMAINSET_POLICY_ROUNDROBIN policies, respectively. The **DOMAINSET_PREF**() policies attempt allocation from the specified domain, but unlike **DOMAINSET_FIXED**() will fall back to other domains to satisfy the request. These policies should be used in preference to **DOMAINSET_FIXED**() to avoid blocking indefinitely on a M_WAITOK request. The **domainset_create**() function takes a partially filled in domainset as a key and returns a valid domainset or NULL. It is critical that consumers not use domainsets that have not been

returned by this function. *domainset* is an immutable type that is shared among all matching keys and must not be modified after return.

The **sysctl_handle_domainset**() function is provided as a convenience for modifying or viewing domainsets that are not accessible via cpuset(2). It is intended for use with sysctl(9).

SEE ALSO

cpuset(1), cpuset(2), cpuset_setdomain(2), bitset(9)

HISTORY

<*sys/domainset.h*> first appeared in FreeBSD 12.0.